



PARIS SCHOOL OF ECONOMICS  
ÉCOLE D'ÉCONOMIE DE PARIS

**WORKING PAPER N° 2011 - 10**

**Adaptive social learning**

**Christoph March**

**JEL Codes: C73, D82, D83**

**Keywords: social learning, informational herding,  
adaptation, analogies, non-Bayesian updating**



**PARIS-JOURDAN SCIENCES ÉCONOMIQUES**

48, Bd JOURDAN – E.N.S. – 75014 PARIS  
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10  
[www.pse.ens.fr](http://www.pse.ens.fr)

# Adaptive Social Learning\*

Christoph MARCH<sup>†</sup>

Preliminary version. Comments are welcome.

This version: February 2011.

## Abstract

This paper investigates the learning foundations of economic models of social learning. We pursue the prevalent idea in economics that rational play is the outcome of a dynamic process of adaptation. Our learning approach offers us the possibility to clarify when and why the prevalent rational (equilibrium) view of social learning is likely to capture observed regularities in the field. In particular it enables us to address the issue of individual and interactive knowledge. We argue that knowledge about the private belief distribution is unlikely to be shared in most social learning contexts. Absent this mutual knowledge, we show that the long-run outcome of the adaptive process favors non-Bayesian rational play.

KEYWORDS: Social Learning; Informational herding; Adaptation; Analogies; Non-Bayesian updating.

JEL CLASSIFICATION: C73, D82, D83.

## 1 Introduction

Does learning by observing others lead to information revelation and efficient social outcomes? To answer this question, early economic models of social learning have analyzed situations in which Bayesian rational individuals are endowed with private signals about a payoff-relevant state of Nature and choose irreversible actions in an exogenous order after having observed their predecessors' actions. Payoff externalities are absent and private signals are discrete but unbiased meaning that the pooled information of individuals reveals the most profitable action. If individuals choose from a continuum of actions and are rewarded according to the proximity of their chosen action to the most profitable action then social learning is efficient (Lee, 1993). The answer to the above question is therefore positive for sufficiently rich action spaces which is obvious since actions perfectly reveal private signals. This conclusion is not entirely satisfactory as there are many economically relevant situations in which individuals cannot fine-tune their actions to their information (Bikhchandani, Hirshleifer, and Welch,

---

\*This paper is a revised version of Chapter II of my PhD thesis (March, 2010) written while I was member of the IMPRS "Uncertainty" at the Max Planck Institute of Economics, Jena and submitted to Technical University Berlin in August 2010. I am particularly indebted to my supervisor Anthony Ziegelmeyer for our discussions and joint work which strongly shaped this paper. I also thank Christophe Chamley, Philippe Jehiel, Dorothea Kübler, Toru Suzuki, Georg Weizsäcker and seminar audiences in Berlin, Jena and Paris for helpful comments and conversations. Financial support from the European Research Council is gratefully acknowledged.

<sup>†</sup>Paris School of Economics, 48 boulevard Jourdan, 75014 Paris, France; Email: MARCH@PSE.ENS.FR

1998; Gale, 1996). In situations where the action space is discrete, an *informational cascade* eventually occurs in which individuals choose actions which do not convey private information and herd on a wrong action with positive probability (Banerjee, 1992; Bikhchandani, Hirshleifer, and Welch, 1992).<sup>1</sup> Assuming a finite number of actions turns the positive answer into a negative answer even though informational cascades do not arise for most atomless distributions of private signals (Chamley, 2004b, chap. 4) and the correct action is chosen asymptotically when private signals are unbounded (Smith and Sørensen, 2000). Indeed, from a social welfare point of view, it matters little whether incorrect herds arise or whether information is fully revealed but extremely slowly. Both phenomena are manifestations of the *self-defeating property* and both phenomena result from the presence of *informational externalities* (Vives, 1996). Learning by observing others is self-defeating for Bayesian rational individuals because the more information has been accumulated in predecessors' decisions the less weight is given to private information in the current decision which in turn yields to a lower increase in public information. Learning by observing others involves informational externalities because Bayesian rational individuals do not take into account the informational effects of their decisions on successors. The self-defeating property and the existence of informational externalities are robust properties of economic models of social learning which imply that the answer to the above question is unequivocally negative whenever actions are discrete (social learning is also inefficient when individuals choose from a continuum of actions but the observation of actions is noisy; see Vives, 1993). This conclusion extends to situations in which individuals choose endogenously the time of their irreversible action (Chamley, 2004a) and, albeit with some qualification, to situations with flexible prices (Vives, 2008).

In addition to being robust, the self-defeating property and the presence of informational externalities are sensible properties which are likely to be consistent with observed regularities in many field environments. Since economic models of social learning possess both properties, they have the potential to deepen our understanding of real-world phenomena like social epidemics. However, several features of the Bayesian rational view of social learning in its current form seem unrealistic or extreme. Most importantly a particularly strong level of sophistication on the part of individuals is assumed which renders them capable of perfectly inferring the degree of information conveyed by any observed action. More precisely any individual is not only assumed to know the structure of everyone's private information, but also the decision model of every other individual including the underlying complete system of beliefs. In other words the social learning context and Bayesian rationality are assumed to be *commonly known*. This assumption has some implications which are unsound and which limit the behavioral relevance of the model. Fully rational individuals correctly take into account that the weight of public information slows down social learning. Therefore even after having observed one thousand identical decisions the average individual is not extremely confident about the appropriateness of the chosen action. When endowed with a sufficiently precise private signal which points in a different direction fully rational individuals overturn the accumulated evidence from many previous actions. According to the overturning principle (Smith and Sørensen, 2000), the belief of the average individual is drastically revised after the observation of such a contrarian action which implies that society eventually learns the truth in rich-enough signal spaces. These conclusions stand in sharp contradiction to the experimental evidence on social learning (Ziegelmeyer, Koessler, Bracht, and Winter, 2010; Weizsäcker, 2010) and they have been criticized by recent models of boundedly rational social

---

<sup>1</sup>Though individuals choose from a continuum of actions in Banerjee (1992), the model shares the properties of a discrete choice model due to degenerate payoffs.

learning (Eyster and Rabin, 2010; Guarino and Jehiel, 2009).

Though we are sympathetic to the bounded rationality approach, we are also convinced that the assumption of fully rational behavior may be justified if it can be proven to arise as the outcome of a dynamic process of learning, or adaptation. In this paper we thus investigate the learning foundations of economic models of social learning. Our premise is that any interactive knowledge in a game must be acquired during repeated play of this or similar games. This perspective enables us to rigorously evaluate the assumptions of common knowledge of both the social learning context and the decision model of each individual.

Our first main result establishes conditions on the adaptive process under which individuals ultimately arrive at rational behavior. These “perfect learning opportunities” (sufficiently many repetitions of a fixed social learning game) are standard in the literature on adaptation and learning in games. Yet, a recurrent issue in this literature is that players will rarely encounter exactly the same game a great number of times. This renders problematic the focus on results in the long run of the adaptive process. One argument in favor of preserving the long run perspective asserts that any sort of learning involves extrapolation across environments which are considered similar (Fudenberg, 2006; Fudenberg and Levine, 1998). Accordingly what matters is how often players encounter similar games. In this paper we not only investigate whether rational play can be justified as the outcome of *some* process of adaptation, but we also analyze the robustness with regard to the conditions which need to be imposed. Indeed our main message is that while perfect learning opportunities might justify the assumption of fully rational behavior, limited learning opportunities such as adaptation taking place across contexts introduce systematic biases into players’ inferences from observed actions. We then show how these biases in turn induce suboptimality of Bayesianism. Hence, in an environment which offers limited learning opportunities individuals who are not Bayesian rational in responding to biased inferences may achieve a higher expected payoff.

The experimental literature on social learning has established systematic deviations of subjects’ behavior from Bayesian rational decision making especially in situations where the private information and the history of public action are conflicting. Indeed subjects generally show a strongly inflated tendency to follow their private information. The results of this paper suggest a reinterpretation of these findings when combined with the idea of *rule rationality* (Aumann, 1997, 2008). Rule rationality asserts that individuals, rather than deriving an optimal strategy in each strategic context separately, behave according to rules which apply and are optimally adapted to a class of contexts. If one accepts that behavior in real-world environments arises as the outcome of adaptation subject to limited learning opportunities our results show that strategies which do not combine public and private information in a Bayesian way may be payoff maximizing. Rule rationality in turn gives a rationale for why subjects may apply such strategies in simple laboratory social learning settings. Consequently, subjects’ behavior in social learning experiments may be an artifact of players’ adjustment to more complex real-world environments and deviations from rational play do not constitute conclusive evidence against rational social learning.

In Section 2 we start with a simple example which illustrates the adaptive process, the distinction between adaptation in a single context and adaptation across contexts, our main result regarding the outcome of the adaptive process in both cases, and how learning across contexts may facilitate non-Bayesian rational strategies. We also employ this example in order to discuss the restrictiveness

of some of our main assumptions.

The framework of social learning is introduced in Section 3 with arbitrary (finite) number of players, binary action and state space and general distribution of private information modeled in the form of private beliefs. We briefly discuss the standard approach which assumes Bayesian rational players and common knowledge of both the structure of the model and Bayesian rationality. We show that these assumptions uniquely characterize the outcome of the social learning game for all but a negligible subset of parameters which implies that rational behavior might in fact be deduced by players.

In Section 4 we set up the adaptive process. We first define the rules governing repeated play of the social learning game including the feedback players receive after each round of playing the entire social learning game (after player  $n$ ). Our most important assumption is that private beliefs of players are never publicly revealed while the state of Nature in general is revealed at the end of the round. We then specify how players learn to *assess* the informational content of sequences of others' choices from feedback of previous rounds. For this purpose we stick closely to the ideas of fictitious play (Brown, 1951): Players assess conditional probabilities of histories (sequence of previous actions) conditional on the realized state of Nature by counting frequencies with which history-state-combinations occurred in previous rounds. Players then respond to these assessed probabilities myopically combining them with private information in a Bayesian way and maximizing the resulting expected payoff.

Section 5 is devoted to studying the outcome of the adaptive process under the standard assumptions that the same fixed social learning game is repeated an arbitrarily large number of times. We show that in this case assessments become correct and behavior approaches rational play. This result is shown to arise as the consequence of a more general one on convergence of fictitious play in dominance-solvable games (Milgrom and Roberts, 1991). A simple implication is that under adaptation in a single context generically the highest expected payoff can only be attained by Bayesian rational individuals.

In Section 6 we turn to the outcome of the adaptive process if the standard assumptions are not satisfied. Primarily we analyze the effect of players learning across multiple social learning games. We restrict ourselves to settings which differ only in the distribution of private information. Under this restriction, in addition to considerations regarding the learning horizon, considerations regarding learning costs, naivety of players, and feedback constraints may provide further justification for learning across settings since players do not possess sufficient information to distinguish settings. We then show that in the long run behavior of players is captured by an *analogy-based expectations equilibrium* (Jehiel, 2005; Jehiel and Koessler, 2008) of the “global” social learning game in which the distribution of private information is randomly determined before the standard social learning game commences. We demonstrate that generically this behavior does not maximize a player's expected payoff in the global social learning game indicating possible benefits for non-Bayesian rational strategies.

We complement the results in the second part of the section by modelling explicitly social learning within a finite number of repetitions. Under this assumption assessments must remain noisy estimates of their true counterparts. We show that noise in assessments causes Bayesian posteriors to in expected terms *underweight* private information. Furthermore we verify again that the combination of Bayesianism and best response is outperformed by other strategies. In particular we indicate that *overweighting of private information* tends to imply a higher expected payoff.

We discuss the paper's relationship to the literature and possible extensions of our results in Section 7. Some concluding remarks are contained in Section 8. The appendix collects various additional

material and most of the proofs.

## 2 A Basic Illustration

In this section, we illustrate the main results of the paper with the help of two simple examples.

### Example 1: Social Learning in a Single Context

Consider a setting where two players, Anna and Bob, face similar investment decisions under uncertainty. Players decide in sequence with Anna deciding first and Bob deciding second after having observed Anna's decision. Payoffs from investing and rejecting are the same for both players. The investment payoff is denoted by the random variable  $\tilde{\theta}$  with possible realizations 0 and 1 which are equally likely.  $\theta$  denotes the true (realized) payoff of the investment. The cost of an investment is  $c = 1/2$  and the payoff to rejecting is zero. Each player has (imperfect) private information about the true value of the investment. More precisely each player  $i \in \{A, B\}$  is endowed with a *private belief*  $b_i$ , the estimated probability that the true payoff of the investment is 1.<sup>2</sup>  $b_A$  and  $b_B$  are conditionally independent realizations of the random variables  $\tilde{b}_A$  and  $\tilde{b}_B$  where  $\tilde{b}_i$ ,  $i = A, B$  is distributed on a finite subset  $B$  of  $(0, 1)$  according to probabilities  $Pr(\tilde{b}_i = b_i | \tilde{\theta} = \theta)$  given the true payoff of the investment  $\theta$ . We assume that private beliefs satisfy the proportional property (Chamley, 2004b, p.31)  $Pr(\tilde{b}_i = b_i | \tilde{\theta} = 1) / Pr(\tilde{b}_i = b_i | \tilde{\theta} = 0) = b_i / (1 - b_i)$  for each  $i$ .<sup>3</sup>

In the following, we first characterize players' behavior predicted by the standard model of rational social learning. It is captured by the unique rationalizable outcome (Bernheim, 1984; Pearce, 1984). Next we show that a (boundedly rational) process of adaptation leads to the same outcome.

### Rationalizable Social Learning

Previous models of rational social learning have relied explicitly or implicitly upon three major assumptions: First, Anna and Bob are *Bayesian rational* which means that they translate all available information into beliefs about the profitability of the investment using the laws of conditional probability and take decisions which maximize expected payoff.<sup>4</sup> Second, Anna's and Bob's decision model is commonly known. We refer to this type of knowledge as (common) *strategic knowledge*. In the social learning context we consider common strategic knowledge is equivalent to common knowledge of Bayesian rationality (see below; in the illustration it is sufficient if Bob knows that Anna is Bayesian rational). Finally, the social learning context, i.e. the payoff structure and the information structure (i.e. conditional distribution of private beliefs) is commonly known (in the illustration it is sufficient

---

<sup>2</sup>Hence  $b_i = Pr(\tilde{\theta} = 1 | \mathcal{I}_i)$  where  $\mathcal{I}_i$  denotes the information of player  $i$ .

<sup>3</sup>The proportional property derives from players whom given private information update a prior in a Bayesian way. A weaker assumption which would allow for an interpretation of private beliefs to be partially shaped by personal opinion is for instance given by  $Pr(\tilde{b}_i = b_i | \tilde{\theta} = 1) / Pr(\tilde{b}_i = b_i | \tilde{\theta} = 0) \leq Pr(\tilde{b}_i = b'_i | \tilde{\theta} = 1) / Pr(\tilde{b}_i = b'_i | \tilde{\theta} = 0)$  if  $b_i < b'_i$ . The results of the example hold if the influence of objective information is sufficiently large.

<sup>4</sup>Hence, players are Bayesian rational in the sense of Tan and Werlang (1988). Bayesian rationality in this sense does NOT comprise knowledge about the rationality of other players. We employ this definition since we desire to separate the two notions. Said differently Bayesian rationality comprises strong instrumental rationality but not necessarily strong cognitive rationality.

if Bob knows that Anna knows). We refer to this type of knowledge as *structural knowledge*.<sup>5</sup> We now derive behavior given these three assumptions.

Anna's belief is straightforwardly given by her private belief  $b_A$ . Therefore Anna's expected payoff from investing is given by  $E[\tilde{\theta} | b_A] - 1/2 = b_A - 1/2$ . Hence, Anna's dominant strategy is to invest if  $b_A > 1/2$ , i.e. if she believes it more likely that the true payoff of the investment is 1, and to reject if  $b_A < 1/2$ .<sup>6</sup> Accordingly, Anna's choice does not depend upon her knowledge about Bob's decision model or the distribution of his private information.

Bob, the second player, in addition to holding a private belief observes Anna's investment decision captured by the history  $h_B \in \{\text{invest, reject}\}$ . Letting  $Pr(\tilde{\theta} = 1 | b_B, h_B)$  denote Bob's belief given his private belief  $b_B$  and the history  $h_B$  it is clear that Bob invests if  $Pr(\tilde{\theta} = 1 | b_B, h_B) > 1/2$  or equivalently if his likelihood ratio

$$\lambda(b_B, h_B) = \frac{Pr(\tilde{\theta} = 1 | b_B, h_B)}{Pr(\tilde{\theta} = 0 | b_B, h_B)} = \frac{b_B}{1 - b_B} * \frac{Pr(\tilde{h}_B = h_B | \tilde{\theta} = 1)}{Pr(\tilde{h}_B = h_B | \tilde{\theta} = 0)}$$

is strictly greater than 1. Since Bob knows that Anna is Bayesian rational (strategic k.) he can "educe" her dominant strategy. Moreover, Bob knows the probabilities that  $b_A > 1/2$  (resp.  $b_A < 1/2$ ) conditional on the true payoff of the investment (structural k.). Since  $Pr(b_A > 1/2 | \tilde{\theta} = 1) > Pr(b_A > 1/2 | \tilde{\theta} = 0)$  and  $Pr(b_A < 1/2 | \tilde{\theta} = 1) < Pr(b_A < 1/2 | \tilde{\theta} = 0)$ <sup>7</sup> Bob's strategy is straightforwardly derived: If Bob's private belief confirms Anna's decision (Anna invests and  $b_B > 1/2$  or Anna rejects and  $b_B < 1/2$ ) then Bob follows his private belief or equivalently imitates Anna's decision. On the other hand if Anna invests and  $b_B < 1/2$  Bob invests (imitates Anna's decision) if

$$(1 - b_B) / b_B < Pr(b_A > 1/2 | \tilde{\theta} = 1) / Pr(b_A > 1/2 | \tilde{\theta} = 0)$$

and rejects (follows his private belief) otherwise. Equivalently if Anna rejects and  $b_B > 1/2$  Bob rejects (imitates Anna) if  $(1 - b_B) / b_B > Pr(b_A < 1/2 | \tilde{\theta} = 1) / Pr(b_A < 1/2 | \tilde{\theta} = 0)$  and invests (follows his private belief) otherwise. This is the dominant strategy for Bob given that Anna has a dominant strategy.

In conclusion Bayesian rationality and strategic and structural common knowledge jointly characterize a unique rationalizable and hence eductively stable (Guesnerie, 1992) outcome which is therefore also the equilibrium outcome.

## Adaptive Social Learning

Contrary to the eductive justification of equilibrium, an adaptive approach does not assume that players are endowed with individual and interactive knowledge before the adjustment process starts. This will be our premise for the remainder of this section. We maintain the assumption that players are Bayesian rational. We assume repeated interactions in the same (fixed) social learning setting and epistemic learning (or beliefs-based learning) which means that players monitor others' decisions across repetitions and extrapolate their future strategies from past decisions using pragmatic boundedly

---

<sup>5</sup>See Brandenburger (1996) for a distinction between strategic and structural uncertainty upon which our distinction between strategic and structural knowledge relies.

<sup>6</sup>If  $b_A = 1/2$ , Anna is indifferent between investing and rejecting. By neglecting this indifference case, we strengthen our conclusion.

<sup>7</sup>See e.g. Milgrom (1981).

rational rules.<sup>8</sup> To simplify exposition, we do not provide much details on the learning-theoretic model. Instead, we simply assume that players do not entertain any strategic repeated play considerations and we abstract from learning costs and discounting. Therefore each time they interact players myopically maximize their expected payoff from the current interaction. The adaptive process is fully exposed in Section 4.

Concretely, let Anna and Bob repeatedly interact in a fixed social learning context (i.e. fixed payoff and information structure). In each round of interaction the investment payoff and both players' private beliefs are drawn anew independently from previous realizations. Furthermore after both Anna and Bob have decided the investment payoff is revealed. Then in each repetition Anna myopically follows her private belief (i.e. invests if  $b_A > 1/2$  and rejects if  $b_A < 1/2$ ) which implies that she invests more often in rounds with an investment payoff of 1. More precisely, the conditional relative frequency of Anna investing given an investment payoff of 1 (0) approaches  $Pr(b_A > 1/2 | \tilde{\theta} = 1)$  ( $Pr(b_A > 1/2 | \tilde{\theta} = 0)$ ). Assume that across repetitions Bob keeps track of the relationship between Anna's choices and the true payoff of the investment. Given sufficiently many rounds Bob will learn the correct conditional relative frequencies and accordingly learn to infer the same information from Anna's decision as he would "educate" when endowed with strategic and structural knowledge. Myopic behavior then induces Bob to eventually play his unique rationalizable strategy.

In conclusion, adaptive learning also leads to the rationalizable outcome.

Our first example suggests that Bayesian rational social learning can be viewed as the long-run outcome of a dynamic process of adjustment. This example is however misleading. Indeed, learning-theoretic models build on the assumptions that the learning horizon is infinite and that there is a large number of players who interact relatively anonymously (to prevent repeated play considerations). The validity of these two assumptions is questionable when learning takes place in a single context.<sup>9</sup> Accordingly, we further investigate the learning-theoretic foundations of Bayesian rational social learning in a second example where players adapt across social learning contexts.

## Example 2: Social Learning in Multiple Contexts

We assume that there are two different social learning settings which differ only in the distributions of private beliefs. Concretely, in setting  $k \in \{H, L\}$  the distributions of Anna's private belief satisfy  $Pr(b_A > 1/2 | \tilde{\theta} = 1, k) = \pi_A^k > 1/2$  and  $1/2 > Pr(b_A > 1/2 | \tilde{\theta} = 0, k) = 1 - \pi_A^k$ <sup>10</sup>, Bob's private belief takes one of two possible values  $0 < 1 - b_B^k < 1/2 < b_B^k < 1$  and we assume that  $b_B^k > \pi_A^k$ .<sup>11</sup> Therefore if Anna and Bob would only interact in one of the settings (or equivalently if Bob would distinguish settings whilst adapting) Bob would eventually play his unique rationalizable strategy and follow his

---

<sup>8</sup>Apparently this process takes place in real time.

<sup>9</sup>Fudenberg and Levine (1998, p. 4) argue that "our presumption that players *do* extrapolate across games they see as similar is an important reason to think that learning models have some relevance to real-world situations."

<sup>10</sup>We assume symmetry only for the ease of exposition.  $\pi_A^k$  then denotes the probability that Anna chooses the more profitable action given the true payoff of the investment. The results straightforwardly extend to the asymmetric case.

<sup>11</sup>We discuss the opposite case at the end of the section.



private belief irrespective of Anna's decision.<sup>12</sup> We assume in addition that  $\pi_A^H > \pi_A^L$  and  $b_B^H > b_B^L$  and accordingly we call  $k = L$  ( $k = H$ ) the low (high) information setting.

As before our premise is that players' strategic and structural knowledge is acquired over time (again it is enough to focus on Bob). Anna and Bob interact repeatedly and Bob keeps track of the same information as before (the relationship between Anna's decision and the realized payoff of the investment). Furthermore as before players myopically maximize the expected payoff of the current interaction. The complication we introduce is that the setting may switch from round to round. Concretely, we assume that each repetition is equally likely to take place in either the low or high information setting.<sup>13</sup> If Bob were able to identify the setting in each round he could simply keep track of Anna's choices in each setting separately. But in general this would require him to possess a significant degree of structural knowledge before the adaptive process starts.<sup>14</sup> In particular, Bob would have to know (or believe) that there exists a correlation between the strength of his private belief and the strength of the information he can derive from Anna's decision. In line with our premise we assume that Bob does not possess this kind of knowledge. We discuss the restrictiveness of this assumption at the end of the section.<sup>15</sup>

The fact that Bob does not distinguish the two settings implies that he learns *across* them. Concretely, in each round Bob uses all the evidence he has accumulated about the relationship between Anna's choice and the investment payoff. Since Anna always follows her private belief and since each setting is equiprobable, Bob eventually infers the following conditional relative frequencies

$$\begin{aligned} Pr(h_B = \text{invest} \mid \tilde{\theta} = 1) &= Pr(h_B = \text{reject} \mid \tilde{\theta} = 0) = \bar{\pi}_A = \frac{\pi_A^L + \pi_A^H}{2}, \text{ and} \\ Pr(h_B = \text{invest} \mid \tilde{\theta} = 0) &= Pr(h_B = \text{reject} \mid \tilde{\theta} = 1) = 1 - \bar{\pi}_A. \end{aligned}$$

Based on these frequencies, Bob adopts the following strategy: In the high information setting, Bob follows his private belief independent of Anna's decision ( $\pi_A^L < \bar{\pi}_A < \pi_A^H < b_B^H$ ); in the low information setting, Bob follows his private belief irrespective of Anna's decision if  $\bar{\pi}_A < b_B^L$  but he imitates Anna's decision irrespective of his private belief if  $\bar{\pi}_A > b_B^L$ . In summary, if  $\bar{\pi}_A > b_B^L$  then Bob suffers an expected loss in the low information setting as he would be better off by following his private belief in this setting as well.

Our second example shows that adaptation across settings may lead Bayesian rational Bob to a suboptimal strategy. Obviously, our assumption that Bob is Bayesian rational was made only for convenience. A more reasonable learning-theoretic framework allows for the adjustment of updating rules since players resort to active experimentation in order to test unusual strategies. We conclude

---

<sup>12</sup>Since  $\lambda(b_B, \text{invest}) = \frac{b_B}{1-b_B} * \frac{\pi_A^k}{1-\pi_A^k} > 1 \Leftrightarrow b_B > 1 - \pi_A^k \Leftrightarrow b_B = b_B^k$  and  $\lambda(b_B, \text{reject}) = \frac{b_B}{1-b_B} * \frac{1-\pi_A^k}{\pi_A^k} < 1 \Leftrightarrow b_B < \pi_A^k \Leftrightarrow b_B = 1 - b_B^k$ .

<sup>13</sup>Our results straightforwardly extend to the non-symmetric case.

<sup>14</sup>By identification of the setting we mean the identification of its information structure. Obviously, social learning contexts differ not only according to their information structure. For instance, the nature of the investment is idiosyncratic to the social learning context. However, unless one is willing to assume that such physical cues are correlated with the information structure our arguments remain valid.

<sup>15</sup>Strictly speaking in this simple example Bob can discriminate the two settings according to the strength of his private belief.

our second example by showing that non-Bayesian rational learning emerges in such an extended framework.

Remember that Bayesian rational Bob invests if his likelihood ratio  $\lambda(b_B, h_B)$  is strictly greater than 1 or equivalently if  $\log\left(\frac{b_B}{1-b_B}\right) + \log\left(\frac{Pr(\tilde{h}_B=h_B|\tilde{\theta}=0)}{Pr(\tilde{s}_h=h_B|\tilde{\theta}=1)}\right) > 0$ . Accordingly, Bayesian rational Bob weights equally his private belief and the information he derives from Anna's decision. Assume alternatively that (non-Bayesian rational) Bob invests provided

$$\beta * \log\left(\frac{b_B}{1-b_B}\right) + \log\left(\frac{Pr(\tilde{h}_B=h_B|\tilde{\theta}=0)}{Pr(\tilde{s}_h=h_B|\tilde{\theta}=1)}\right) > 0$$

where  $\beta > 0$  captures the weight of the private belief relative to the weight of the information inferred from Anna's decision. In this case Bob's strategy is characterized as follows: In each setting  $k \in \{L, H\}$ , there exists  $\beta_k^*$  such that Bob follows his private belief irrespective of Anna's decision provided  $\beta > \beta_k^*$  and Bob imitates Anna's decision irrespective of his private belief provided  $\beta < \beta_k^*$ .<sup>16</sup>  $\beta_L^* > \beta_H^*$  implies that Bob follows his private belief irrespective of Anna's decision in either setting provided  $\beta > \beta_L^*$ . Therefore, a non-Bayesian rational Bob whose relative weight on his private belief is strictly greater than  $\beta_L^*$  achieves higher fitness than a Bayesian rational Bob if  $\beta_L^* > 1$ .

In conclusion, if Bob cannot distinguish between the two settings then he is not able to infer the correct magnitude of information that Anna's decision conveys in each setting. In other words, the complete resolution of uncertainty (strategic and structural) is not possible in the presence of large structural uncertainty.<sup>17</sup> It may then be beneficial for Bob not to form his belief in a Bayesian way. We conjecture that our result holds for most adaptive processes though a proof of this assertion is left for future research.

## Discussion

In our second example, Bob has the possibility to straightforwardly discriminate the two settings according to the strength of his private belief. This is not possible in an extended framework. Assume that Bob can be weakly ( $b_B \in \{1 - b_B^L, b_B^L\}$ ) or well informed ( $b_B \in \{1 - b_B^H, b_B^H\}$ ) in *both* settings but that he is more likely to be weakly informed in the low information setting. Concretely, assume that in the low (high) information setting the probability that Bob is weakly (well) informed is given by  $\alpha > 1/2$ . In this case it is optimal for Bob to follow an extreme private belief  $b_B \in \{1 - b_B^H, b_B^H\}$  since  $\pi_A^L < b_B^L < \pi_A^H < b_B^H$  and it is optimal for him to follow a less extreme private belief  $b_B \in \{1 - b_B^L, b_B^L\}$  provided  $b_B^L > \alpha \pi_A^L + (1 - \alpha) \pi_A^H$ .<sup>18</sup> In particular  $\alpha \pi_A^L + (1 - \alpha) \pi_A^H < \bar{\pi}_A$  since  $\alpha > 1/2$  and there exists a set of parameters such that  $\alpha \pi_A^L + (1 - \alpha) \pi_A^H < b_B^L < \bar{\pi}_A$ , i.e. such that Bob given a less extreme private belief imitates Anna's decision although he should optimally follow his private belief. Hence,

---

<sup>16</sup>  $\beta_k^* = \log(\bar{\pi}_A / (1 - \bar{\pi}_A)) / \log(b_B^k / (1 - b_B^k))$ .

<sup>17</sup> Large structural uncertainty prevails in field environments since the generating process of players' private beliefs is rarely commonly known (see Dekel and Gul, 1997, p. 101). This being said, real-world situations with low structural uncertainty also exist. Assume for instance that the payoff of the investment is determined by the uncertain amount of oil in some tract and that players' private beliefs result from each of them taking a soil sample (Hendricks and Kovenock, 1989). Published experiments provide a thorough understanding of both the prior likelihood of oil and the distribution of samples as a function of the oil in the tract. In this situation, (common) knowledge of the distribution of private beliefs is easily justified.

<sup>18</sup>  $Pr(\tilde{\theta} = 1 | b_B^L, h_B) = \frac{\alpha b_B^L Pr(h_B|\tilde{\theta}=1, L) + (1-\alpha) b_B^L Pr(h_B|\tilde{\theta}=1, H)}{\alpha b_B^L Pr(h_B|\tilde{\theta}=1, L) + (1-\alpha) b_B^L Pr(h_B|\tilde{\theta}=1, H) + \alpha (1-b_B^L) Pr(h_B|\tilde{\theta}=0, L) + (1-\alpha) (1-b_B^L) Pr(h_B|\tilde{\theta}=0, H)}$  and equivalently for  $Pr(\tilde{\theta} = 1 | 1 - b_B^L, h_B)$ .

we reproduce the result that learning across settings may lead to suboptimal behavior for Bayesian rational players.

Bayesian rational Bob could still insist on distinguishing between the two settings and he would eventually arrive at the optimal strategy.<sup>19</sup> In assuming that Bob does not condition on his own quality of information we thus assume that he learns *imperfectly* from feedback. Hence, Bob is *naive* akin to players in Esponda (2008a). Yet, we doubt that perfect learning from feedback is always a sensible assumption. Miettinen (2010) argues that own types may not be remembered over time, in particular if they are not part of inherent personal characteristics. Additionally, memory, information processing and feedback constraints may restrain the number of categories a player can separate while learning.

Finally, if players rely on the adaptive process to assess both their own information and uncertainty then Bayesian rationality eventually yields the strategy with the highest fitness. We find this argument unconvincing. Social learning occurs in situations where informed players (though imperfectly) try to learn from observing others' decisions.

### 3 Preliminaries

#### 3.1 The Social Learning Stage Game

A finite number of players  $i = 1, 2, \dots, n$  choose an action, in that exogenous order, from the set  $A = \{0, 1\}$ . Each player's payoff depends on the realization of an underlying state of Nature and the chosen action. The state of Nature is given by the random variable  $\tilde{\theta}$  distributed on  $\Theta = \{0, 1\}$ , over which players share a common prior belief. Without loss of generality, the prior is assumed to be flat, with both states equally likely. Players' payoffs are given by the mapping  $u : A \times \Theta \rightarrow \mathbb{R}$  where  $u(1, \theta) = \theta - \frac{1}{2}$  and  $u(0, \theta) = 0$  for each  $\theta \in \Theta$ . In the following, player  $i$ 's action  $a_i = 1$  (resp.  $a_i = 0$ ) is sometimes referred to as "invest" (resp. "reject") and the cost of the investment is set equal to  $\frac{1}{2}$  merely to simplify the exposition (our results straightforwardly extend to any cost in the interval  $(0, 1)$ ). In a similar vein, both the underlying state of Nature and the action set are binary to simplify the exposition (our results extend to any finite number of actions and states but at significant algebraic cost).

The realized state  $\theta$  is unknown and each player is endowed with some (imperfect) private information about the realized state. Before any action is taken, player  $i$ 's imperfect knowledge about the realized state is called her *private belief* and is denoted by  $b(\tilde{s}_i, \emptyset)$ . This endowment of player  $i$  is a probability estimate of the state of Nature which can be interpreted as resulting from the prior probability of the state and player  $i$ 's private signal  $\tilde{s}_i$ . In the following, we often identify  $\tilde{s}_i$  (resp.  $s_i$ ) with the private belief (resp. realization of the private belief)  $b(\tilde{s}_i, \emptyset)$  (resp.  $b(s_i, \emptyset)$ ). Conditional on the realized state,  $(\tilde{s}_i)_{i=1}^n$  is an i.i.d. sequence generated according to the c.d.f.  $G_\theta(s)$ .  $G_0$  and  $G_1$  satisfy the standard assumptions meaning that they have common support whose convex hull is given by  $[\underline{b}, \bar{b}]$  and their Radon-Nikodym derivative is such that  $\frac{dG_1}{dG_0}(s) = \frac{s}{1-s}$ . We assume that  $\underline{b} > 0$

---

<sup>19</sup>To see this note that across the subset of repetitions such that  $b_B \in \{1 - b_B^H, b_B^H\}$  ( $b_B \in \{1 - b_B^L, b_B^L\}$ ) the relative frequency with which Anna invests when the true payoff of the investment is 1 and the relative frequency with which she rejects when the true payoff of the investment is 0 each approaches the true conditional probability  $\alpha \pi_A^H + (1 - \alpha) \pi_A^L$  ( $\alpha \pi_A^L + (1 - \alpha) \pi_A^H$ ) conditional on Bob holding a high (low) quality of information.

and  $\bar{b} < 1$  i.e. private beliefs are bounded. Note that the property on the Radon-Nikodym derivative implies that  $G_0$  and  $G_1$  satisfy the property of first-order stochastic dominance:  $G_1(s) < G_0(s)$  for each  $s \in (\underline{b}, \bar{b})$ .

Though player  $i$  cannot observe the private belief of any other player, she observes the complete history  $h_i = (a_1, \dots, a_{i-1}) \in H_i = A^{i-1}$  of previous actions with  $h_1 \equiv \emptyset$ . In fact, we assume that history  $h_i$  is observed by all players  $i, i+1, \dots, n$  and that this knowledge is common to all players.  $H_{n+1} = A^n$  denotes the set of complete histories with element  $h_{n+1} = (a_1, \dots, a_n)$  and  $H = \bigcup_{i=1}^n H_i$ . Given a sequence of actions  $(a_1, \dots, a_{i-1})$ , the probability estimate of the state of Nature that is based solely on the public information is called the *public belief* and is given by  $b(\emptyset, h_i) = \Pr(\tilde{\theta} = 1 \mid h_i) = \Pr(h_i \mid \tilde{\theta} = 1) / (\Pr(h_i \mid \tilde{\theta} = 1) + \Pr(h_i \mid \tilde{\theta} = 0))$  with  $b(\emptyset, h_1) = 1/2$ .

We denote by  $\langle n, A, u, \Theta, (G_0, G_1) \rangle$  the *social learning game*. Smith and Sørensen (2000) study a non-straightforward generalization of the social learning game where players have heterogeneous preferences which are private information and some of them have state independent preferences with a single dominant action. In subsection 6.2, we discuss how our results extend to this generalized social learning game.

We conclude this subsection with a definition (Smith and Sørensen, 2000).

**Definition 1.** A property is **generic** or **robust** if it holds for an open and dense subset of parameters of the social learning game.

### 3.2 Bayesian Rational Play

We now define rational play in the social learning stage game. Without loss of generality we model players' behavior in the social learning game by *behavioral strategies*  $\sigma_i : [\underline{b}, \bar{b}] \times H_i \rightarrow \Delta(A)$ ,  $i = 1, \dots, n$ , where  $\sigma_i(s_i, h_i)$  denotes player  $i$ 's probability of investment given realized private belief  $s_i$  and realized history  $h_i$ . We denote by  $\Sigma_i$  the strategy set of player  $i$ . Furthermore as is standard we let  $\Sigma = \times_{i=1}^n \Sigma_i$  denote the set of strategy profiles and  $\Sigma_{-i} = \times_{j \neq i} \Sigma_j$  denote the set of strategy profiles of player  $i$ 's opponents.

In order to derive rational predictions for the social learning game we rely on the following three assumptions. First, players are Bayesian rational in the sense of Tan and Werlang (1988, Axioms (B) and (B.R.)). This means that players are Bayesian and make decisions which maximize their expected payoffs. Second, Bayesian rationality is commonly known. We refer to this as *strategic (common) knowledge*. Finally, the structure of the social learning game, more precisely its information structure, its payoff structure, and players' preferences, are common knowledge among players. We refer to this as *structural (common) knowledge*.

As made clear by Tan and Werlang (1988) the premises above restrict players to iteratively undominated strategies or equivalently correlated-rationalizable strategies (see also Bernheim, 1984; Pearce, 1984; Perea, 2001). In Proposition 1 below we characterize these strategies in terms of Bayesian consistent beliefs  $b_i : [\underline{b}, \bar{b}] \times H_i \rightarrow [0, 1]$  where

$$b_i(s_i, h_i) = \Pr(\tilde{\theta} = 1 \mid b(\tilde{s}_i, \emptyset) = s_i, \tilde{h}_i = h_i)$$

and associated sequential best responses. For this purpose player  $i$ 's *ex-ante expected payoff* of strategy

$\sigma_i$  given strategy profile  $\sigma_{-i}$  of the opponents is given by

$$U_i(\sigma_i, \sigma_{-i}) = \frac{1}{2} \sum_{\theta \in \Theta} \sum_{h_i \in H_i} Pr(h_i | \tilde{\theta} = \theta, \sigma) \int_{\underline{b}}^{\bar{b}} \sum_{a \in A} \sigma_i(a | s, h_i) u(a, \theta) dG_\theta(s). \quad (1)$$

Strategy  $\sigma_i \in \Sigma_i$  is *strictly dominated* if there exists  $\sigma'_i \in \Sigma_i$  such that  $U_i(\sigma'_i, \sigma_{-i}) > U_i(\sigma_i, \sigma_{-i})$  for each  $\sigma_{-i} \in \Sigma_{-i}$ . Iterated elimination of strictly dominated strategies is defined accordingly. We say that a strategy profile is *iteratively undominated* if any of its components survives the elimination procedure.

**Proposition 1.** *To any iteratively undominated strategy profile  $\sigma$  there exists a belief system  $\{b_i\}_{i=1}^n$  such that (i) beliefs are formed according to **Bayes' rule**, i.e. for each  $i = 1, \dots, n$*

$$b_i^*(s_i, h_i) = \frac{s_i Pr(h_i | \tilde{\theta} = 1, \sigma^*)}{s_i Pr(h_i | \tilde{\theta} = 1, \sigma^*) + (1 - s_i) Pr(h_i | \tilde{\theta} = 0, \sigma^*)}$$

*provided  $Pr(h_i | \tilde{\theta} = \theta, \sigma^*) > 0$  for some  $\theta \in \Theta$  where  $Pr(h_i | \tilde{\theta} = \theta, \sigma^*) = \prod_{j < i} \int_{\underline{b}}^{\bar{b}} \sigma_j^*(a_j | s_j, h_j) dG_\theta(s_j)$ ,  $a_j = h_i(j)$  and  $h_j \subset h_i$ , and (ii) behavioral strategies are **sequentially rational**, i.e. for each  $i = 1, \dots, n$*

$$\sigma_i^*(s_i, h_i) = \begin{cases} 1 & \text{if } b_i^*(s_i, h_i) > \frac{1}{2} \\ 0 & \text{if } b_i^*(s_i, h_i) < \frac{1}{2} \end{cases}.$$

*Barring non-genericities all iteratively undominated strategy profiles yield the same unique outcome of the game.*

The proof of this proposition is relegated to the appendix. In the following we denote by  $\sigma^*$  the strategy profile satisfying properties (i) and (ii) of the proposition. Note that iterated dominance does not restrict players' behavior in case of a tie, i.e. whenever  $b_i^*(s_i, h_i) = 1/2$  or equivalently  $b_i(s_i, \emptyset) = 1 - b(\emptyset, h_i)$ . However, as shown in the proof the occurrence of ties with strictly positive probability is a non-generic property of a social learning setting. Therefore w.l.o.g. we assume henceforth that the social learning setting does not allow for ties which implies that there is no need to commit to a specific tie-breaking rule. Absent ties iterated dominance yields a unique outcome which is therefore also the unique Bayesian equilibrium outcome.

## 4 The Adaptive Process

The rational benchmark derived in the previous section presupposes a substantial amount of information on the part of players both about the environment and about others' degree of rationality. In this section we describe a learning process according to which myopic players who do not possess *a priori* all necessary information might learn to play rationally the social learning game. First, we discuss the properties of the learning environment. Second, we detail the rules guiding players' learning by linking our approach with fictitious play. Finally, we discuss how our approach relates to other learning approaches.

### 4.1 Environment

We consider an *extended social learning game*  $\langle n, A, u, \Theta, (G_0, G_1), R \rangle$  where players play the social learning game repeatedly in rounds  $r = 1, 2, \dots, R$ . For the sake of clarity, we assume that the social

learning game is known by all players (this assumption is relaxed in section 5.2). Though we assume mutual knowledge of the social learning game, we refrain from assuming any higher-order interactive knowledge. Since the elements of the social learning game, most importantly the distribution of private beliefs, remain unchanged across rounds we call this environment *stable*.

The following additional assumptions govern repeated play of the social learning game. First, in each round a player's position in the sequence is determined randomly and independently from previous rounds according to a uniform distribution. Hence in each round any player is equally likely to act in any period  $i \in \{1, \dots, n\}$ . Second, the state of Nature as well as every player's private belief is drawn anew each round independently from respective realizations in previous rounds. With this assumption we rule out that learning *about the state* confounds strategic learning as our focus is on the latter.<sup>20</sup> Finally, payoffs are realized immediately at the end of each round.

In order to give players the chance to learn from repeated play players receive feedback at the end of each round of play. Our most important assumption which will be kept throughout is that private beliefs are never publicly revealed (notice that obviously a player may always recall her own private belief realization). Hence, at the end of a round a player may learn at most the realized state of Nature and the actions of all other players. In a first step we will assume that in each round every player indeed receives this maximal feedback. Formally, for each  $i = 1, \dots, n$  and each  $r = 1, \dots, R$  player  $i$ 's feedback in round  $r$  is given by  $y_i(r) = (h_{n+1}(r), \theta(r))$  where  $h_{n+1}(r)$  denotes the complete history of choices and  $\theta(r)$  the realized state of Nature in round  $r$ . We discuss how adaptation is affected if feedback is further constrained in Section 7.

## 4.2 Mechanism

We now detail the rules governing players' learning. We stick closely to the ideas underlying the concept of *fictitious play* (Brown, 1951).

Denote by  $\mathbf{y}(r)$  the vector of feedback players receive at the end of round  $r$ . In each round  $r$ ,  $\mathbf{y}(r) \in Y = A^n \times \Theta$ . Let  $\zeta_r = (\mathbf{y}(1), \dots, \mathbf{y}(r-1))$  denote the collection of feedback at the beginning of round  $r$ . Clearly,  $\zeta_r \in \mathcal{Z}_r = Y^{r-1}$  where  $\mathcal{Z}_1 = \emptyset$ . We call  $\zeta_r$  a realized *learning path* in round  $r$ . Accordingly  $\zeta_\infty \in \mathcal{Z}_\infty = Y^\infty$  denotes an infinite learning path.

Non-revelation of others' private beliefs rules out the opportunity for players to learn others' complete behavioral strategies  $\sigma : [b, \bar{b}] \times H \rightarrow [0, 1]$ <sup>21</sup> given that we abstract entirely from eductive reasoning at this point. Yet, in the absence of payoff externalities players care about others' strategies only insofar as to be able to draw inferences from observed choices about the realized state of Nature. More precisely in order to learn from observing others' choices it is sufficient for players to know conditional probabilities of histories  $Pr(h \mid \tilde{\theta} = \theta)$  conditional on the realized state of Nature.

Let  $y(r)$  (resp.  $\zeta_r$ ) denote feedback (resp. learning paths) for a fixed representative player. As in fictitious play we assume that players learn via a frequentist approach. That is players keep track of

---

<sup>20</sup>For instance if to the contrary the state were fixed once and for all in round 1, a player receiving new private information each round would eventually learn this state as  $R \rightarrow \infty$ . More generally any correlation of states across rounds confounds learning about the state and about other players' strategic behavior.

<sup>21</sup>Notice that contrary to the one-shot variant in the extended social learning game the random assignment of positions requires every player's strategy to determine behavior at **all** possible histories.

frequencies of histories conditional on the realized state. Formally, let

$$\kappa(h_i, \theta | \zeta_r) = |\{1 \leq \rho \leq r : \{h_i, \theta\} \subseteq y(\rho)\}| \quad (2)$$

for  $i = 1, \dots, n$  denote the number of times across rounds  $\rho = 1, \dots, r$  along learning path  $\zeta_r$  that history  $h_i$  occurred when the observed state was  $\theta$ . Players then use these frequencies to form estimates of the state-contingent probabilities of histories. Formally for  $i = 1, \dots, n$  we define *assessments*  $\hat{\varphi}_i : \Theta \times \bigcup_{r=1}^{\infty} \mathcal{Z}_r \rightarrow \Delta(H_i)$  recursively via  $\hat{\varphi}_1(h_1 | \theta; \zeta_r) = 1$  for each  $r = 1, \dots, R$ ,  $\zeta_r \in \mathcal{Z}_r$  and  $\theta \in \Theta$  and

$$\hat{\varphi}_i(h_i | \theta; \zeta_r) = \frac{\kappa(h_i, \theta | \zeta_r) + \epsilon * \hat{\varphi}_{i-1}(h_{i-1} \subset h_i | \theta; \zeta_r)}{\sum_{h'_i \in H_i} [\kappa(h'_i, \theta | \zeta_r) + \epsilon * \hat{\varphi}_{i-1}(h_{i-1} \subset h'_i | \theta; \zeta_r)]} \quad (3)$$

for each  $i = 2, \dots, n$ , each  $h_i \in H_i$ , and each  $r = 1, \dots, R$ ,  $\zeta_r \in \mathcal{Z}_r$ , and  $\theta \in \Theta$ . Notice that contrary to fictitious play we do not assume that players have initial weights  $\kappa(h, \theta | \emptyset) > 0$  for each  $h$  and  $\theta$  since we also study the process with finite repetitions where the influence of such weights is not negligible. We deal with the entailing problem of degeneracy by relying on the idea of trembles: as long as a player cannot rely on accumulated evidence she attaches a small probability  $\epsilon > 0$  to each observation she cannot rule out a priori. We then study the limit as  $\epsilon \rightarrow 0$ . Accordingly players attach probability zero to histories never observed before. On the other hand beliefs at such histories are well-defined: Choices in a history observed for the first time are treated as uninformative about the state.<sup>22</sup>

In defining players' responses to assessments we return to the fictitious play approach. In particular we assume that players in any round take into account their current observations in a myopic way. In other words players do not engage in strategic repeated play considerations, an assumption justified for instance in large population models. To determine myopic response to observations fix an assessment  $\hat{\varphi}$  and denote by  $\hat{\sigma} : \bigcup_{r=1}^{\infty} \mathcal{Z}_r \rightarrow \Sigma$  a *strategic response*.  $\hat{\sigma}$  is *myopic Bayes-rational* iff

$$\hat{\sigma}(s, h | \zeta_r) = \begin{cases} 1 & \text{if } b(s, \emptyset) > 1 - \hat{b}(\emptyset, h | \zeta_r) \\ 0 & \text{if } b(s, \emptyset) < 1 - \hat{b}(\emptyset, h | \zeta_r) \end{cases} \quad (4)$$

where

$$\hat{b}(\emptyset, h | \zeta_r) = \frac{\hat{\varphi}(h | 1; \zeta_r)}{\hat{\varphi}(h | 1; \zeta_r) + \hat{\varphi}(h | 0; \zeta_r)} \quad (5)$$

is the *assessed public belief* at history  $h$  along learning path  $\zeta_r$ . Assuming myopic Bayes-rational responses also means that we abstract from learning costs i.e. players maximize the undiscounted sum of expected payoffs in each repetition. We do this to give best chances to the emergence of the rationalizable outcome. In subsection 6.2 we return briefly to the issue of learning costs.

### 4.3 Limit Outcome

Ultimately, we are interested in assessments and strategic responses in the adaptive process after a sufficient number of repetitions. More precisely we wish to study the relationship to their counterparts

---

<sup>22</sup>Sometimes it is more convenient to work with conditional choice probabilities  $Pr(a_i | h_i, \theta)$ . We thus define in a slight abuse of notation  $\hat{\varphi}(a_i | h_i, \theta; \zeta_r) = \hat{\varphi}_{i+1}((h_i, a_i) | \theta; \zeta_r) / \hat{\varphi}_i(h_i | \theta; \zeta_r)$ . This incorporates the assumption that players' assessments obey independence across periods which is weak given the structure of the social learning stage game considered. It is straightforward to restate equation (3) in terms of these assessments using choice frequencies  $\kappa(a, h, \theta | \zeta_r)$  defined in the obvious way.

in the (unique) outcome of rational play. Accordingly it is necessary to define ideas of closeness for both assessments and strategies.

We let  $\sigma^*$  denote some iteratively undominated strategy profile with associated assessments  $\varphi^* = (\varphi_i^*)$  where  $\varphi_i^*(h_i | \theta) = \Pr(h_i | \tilde{\theta} = \theta, \sigma^*)$  for  $i = 1, \dots, n$ ,  $h_i \in H_i$  and  $\theta \in \Theta$ .<sup>23</sup> We then start by noting that assessments can be interpreted as a matrix, i.e.  $\varphi \in [0, 1]^{(2^n-1) \times 2}$ . It is thus natural to measure distance of assessments using the metric on the associated space of matrices induced by some matrix norm where w.l.o.g. we take the max norm. With regard to strategies we follow Jackson and Kalai's (1997) notion of  $\epsilon$ -like play. In the appendix (lemma D.1 in appendix D) we show that both notions of closeness are mutually dependent.

**Definition 2.** Fix player  $i$  and let  $\epsilon > 0$ . On learning path  $\zeta_r$  in round  $r$  of the adaptive process

- assessments  $\hat{\varphi}_i(\cdot | \cdot; \zeta_r)$  are  $\epsilon$ -close to rational assessments  $\varphi_i^*$  if  $\max \{|\hat{\varphi}_i(h_i | \theta; \zeta_r) - \varphi_i^*(h_i | \theta)| : h_i \in H_i, \theta \in \Theta\} < \epsilon$ ,
- strategic response  $\hat{\sigma}_i$  plays  $\epsilon$ -like  $\sigma_i^*$  at  $h_i \in H_i$  if there exists  $B_\epsilon \subseteq [\underline{b}, \bar{b}]$  such that  $[G_0(B_\epsilon) + G_1(B_\epsilon)]/2 > 1 - \epsilon$  and  $|\hat{\sigma}_i(s_i, h_i) - \sigma_i^*(s_i, h_i)| < \epsilon$  for each  $s_i \in B_\epsilon$ .

Given the objective distributions over the state of Nature and private beliefs the adaptive process defined above gives rise to a probability distribution over learning paths  $\tilde{\zeta}_R$  where  $R \in \mathbb{N} \cup \{\infty\}$ . We will denote this distribution by  $\mathbf{P}$ . All our probabilistic results are with respect to this distribution.

We conclude this subsection with a definition.

**Definition 3.** If  $R \rightarrow \infty$  (resp.  $R < \infty$ ) then the learning horizon is said to be *infinite* (resp. *finite*).

#### 4.4 Relationship to Other Approaches

As discussed in the introduction we place ourselves into a huge literature on learning in games. In particular the adaptive process set up above takes up ideas from the “fictitious play” model (Brown, 1951) and adjusts them to a social learning setting and the associated likely restrictions upon players feedback after each round of play. Our model is thus very specific in its description of how players reach decisions given their experience from previous rounds of play. More general models of learning which are significantly less specific about players’ decision processes have been developed for instance by Milgrom and Roberts (1991) for normal-form games and Fudenberg and Kreps (1995) for extensive-form games. In the lemma below (proven in appendix D) and in appendix A we establish that the adaptive process defined above can be seen as a special case of either of these more general approaches. The convergence result for the standard setting (Proposition 2 in the next section) can thus be extended to a general class of adaptive processes. In particular dominance-solvability of the game (in the generic case) implies that our result follows straightforwardly from the lemma below, Proposition 1 and Theorem 5 in Milgrom and Roberts (1991). Still, we consider our focus on an explicit adaptive process useful as it permits us to take into account specificities of the social learning game at hand. Indeed it is not at all clear that the process with limited feedback on others’ strategies should satisfy *consistency with adaptive learning* as defined by Milgrom and Roberts (1991).

---

<sup>23</sup>Notice that  $\varphi^*$  is the same for any iteratively undominated strategy profile  $\sigma^*$  since we rule out ties.



**Lemma 1.** *With infinite learning horizon the sequence of strategy profiles  $\hat{\sigma}(\zeta_r)$  generated by the adaptive process is almost surely consistent with adaptive learning as defined by Milgrom and Roberts (1991).*

## 5 Perfect Learning Opportunities

This section presents the first of our key results: If the environment is stable and the learning horizon is infinite then the learning process eventually approaches the equilibrium outcome. We also argue that this does not depend on knowledge about the primitives of the model. A simple consequence is that under such “perfect” learning opportunities, Bayes-rational responses to posteriors maximize the (ex-ante) expected payoff.

### 5.1 Learning Conditions for Rational Play

**Proposition 2.** *Assume perfect learning opportunities, i.e. a stable learning environment and an infinite learning horizon. Along almost any learning path  $\zeta_\infty$  and for each  $\epsilon > 0$  strategic responses eventually play  $\epsilon$ -like any iteratively undominated strategy at all histories occurring with strictly positive (uniquely defined) probability under such strategies. Equivalently along almost any learning path  $\zeta_\infty$  and for each  $\epsilon > 0$  assessments are eventually  $\epsilon$ -close to rational assessments.*

While the proof of this proposition is a direct consequence of Proposition 1 and Lemma 1 and Theorem 5 in Milgrom and Roberts (1991) we re-prove it here in the specific form stated above.

*Proof.* The proof uses a similar inductive argument as employed already in the proof of Proposition 1: In period 1 no inferences from others can be drawn. Therefore assessments for the first period are equal to rational assessments by definition and play of the dominant strategy follows straightforwardly from the assumption of myopic Bayesian rational strategic responses. A version of the law of large numbers (SLLNCE, see e.g. Walk, 2008, and the references therein) then implies that assessments in the second period converge to their rational counterparts. In the appendix (Lemma D.1) we establish that given Bayesian rationality this is sufficient to guarantee that strategies eventually play  $\epsilon$ -like any iteratively undominated strategy at all histories occurring with strictly positive probability under this strategy (recall that besides in non-generic settings probabilities of all histories and behavior at histories occurring with strictly positive probability are uniquely defined by iterated dominance). This argument can be inductively extended to all positions  $i > 1$  using again the law of large of numbers and the conversion established in Lemma D.1 that if chosen strategies play  $\epsilon$ -like iteratively undominated strategies in all periods  $j < i$  then assessments in period  $i$  derived from these strategies must be close to rational assessments.  $\square$

Perfect learning opportunities correspond to the necessary<sup>24</sup> and sufficient conditions that permit beliefs and choices to approach their rational counterparts during the adaptive process. The following result is a straightforward consequence of the proposition above.

---

<sup>24</sup>Necessity is shown in the next section.

**Corollary.** *With perfect learning opportunities, for each period  $i = 1, \dots, n$  the Bayes-rational strategy  $\sigma_i^*$  almost surely eventually maximizes the (ex-ante) expected payoff and uniquely so in some social learning games.*

## 5.2 Ex-ante Knowledge

So far we have assumed that the primitives of the game are known by the players *a priori*. Under perfect learning opportunities this assumption seems overly restrictive.<sup>25</sup> Indeed it can be argued that with perfect learning opportunities even completely uninformed players may learn to play rationalizable in a social learning game.

Recall that the structure of the social learning game comprises the state space and the prior, the distribution of each player's private information, every player's action set, and the utility function. We first notice that the adaptive process is such that a player cannot and need not learn private belief distributions of her opponents. Furthermore we maintain the assumption that a player is informed about her own action set *a priori* to abstract from learning about strategy spaces (see e.g. Fudenberg and Levine, 1998, Section 4.9). In line with the adaptive process we assume that players use frequentist models. That is in each round elements (states, private belief realization, ...) which have previously occurred are attached  $1-\epsilon$  times their relative frequency while mass  $\epsilon$  is attached to previously unobserved elements<sup>26</sup>. With this assumption it is immediately clear that players can fully explore the true state and private belief space. Second, learning the associated (conditional) distributions is the classical statistical problem of learning from i.i.d. data by frequentist statisticians. Formally, our adaptive process is easily extended in this direction defining frequencies  $\kappa(\theta \mid \zeta_r)$  and  $\kappa(B, \theta \mid \zeta_r)$  for  $B \subset [0, 1]$  and assessments  $\hat{\phi}(\theta \mid \zeta_r) = (1-\epsilon) * \kappa(\theta \mid \zeta_r)/r$  and  $\hat{\phi}(B \mid \theta; \zeta_r) = (1-\epsilon) * \kappa(B, \theta \mid \zeta_r)/\kappa(\theta \mid \zeta_r)$  and noting that a player may always use (assessed) conditional probabilities to update her (assessed) prior given her realized private belief in a Bayesian way.<sup>27</sup> The law of large numbers then guarantees<sup>28</sup> that assessed private beliefs approach their true counterparts. Third, it is easy to show using the sequential structure of the game that any action which is played with strictly positive probability at some history in the rationalizable outcome will be played. Therefore players will explore others' action spaces. Finally, the infinite learning horizon permits players to experiment sufficiently often in order to explore the structure of their utility function.

---

<sup>25</sup>We discuss consequences of a lack of ex-ante knowledge in the absence of perfect learning opportunities in Section 6.

<sup>26</sup>We recognize that this is a vague argument which omits specifying formally the prior assumptions of players. Yet, it reflects the idea that players' learning cannot be overthrown by unanticipated realizations. Formally it may be modelled by players having non-degenerate priors on appropriate underlying spaces, but choice of the latter remains an open question.

<sup>27</sup>Formally  $\hat{b}(s, \emptyset \mid \zeta_r) = \frac{\hat{\phi}(1; \zeta_r) \hat{\phi}(s|1; \zeta_r)}{\hat{\phi}(1; \zeta_r) \hat{\phi}(s|1; \zeta_r) + \hat{\phi}(0; \zeta_r) \hat{\phi}(s|0; \zeta_r)}$ .

<sup>28</sup>In a recent paper, Al-Najjar (2009) shows that successfully learning from i.i.d. data might be hard in some settings. More precisely if the underlying space is discrete infinite and the  $\sigma$ -algebra of all subsets has to be learned, learning the correct probability is impossible. Since in our case the state space is binary and the private belief space bounded, his results do not apply to this setting provided we consider well-behaved subsets  $B$ , e.g.  $B = [0, b]$  for some  $0 < b < 1$ .

## 6 Limited Learning Opportunities

### 6.1 Adaptation across Games

Our convergence result (Proposition 2) requires that players play the same social learning game a great (infinite) number of times. In the real world this will rarely be the case. It has however been argued that “any sort of learning involves extrapolation from past observations to settings that are deemed (implicitly or explicitly) to be similar, so what matters is how often agents have played ‘similar’ games” (Fudenberg, 2006, p. 701). As discussed in the introduction this idea is prevalent in economics but until recently has rarely been modeled explicitly. We now extend our setting in order to investigate which strategic behavior emerges in a social learning game if adaptation takes place across contexts. Throughout this subsection we maintain the assumption that the learning horizon is infinite.

#### Extension of the Adaptive Process

We consider an (*extended*) *global social learning game*  $\langle n, A, u, \Theta, (G_0^k, G_1^k)_{k=1}^K, \pi(\cdot, R) \rangle$ . For each  $k = 1, \dots, K$ ,  $\langle n, A, u, \Theta, (G_0^k, G_1^k) \rangle$  is a standard social learning game. The global game differs from this game only by an additional move of Nature taking place before the start of the standard game which determines randomly the relevant distribution of private beliefs  $(G_0^k, G_1^k)$ . The random draw is according to probability vector  $\pi \in [0, 1]^{K \cdot 2^9}$ , independent of all random draws within the standard game and unobserved by players. Furthermore in the extended version of the game players repeatedly play the global game where in each repetition (round) the relevant private belief distribution is drawn anew and independently from the distribution in previous rounds.

As before we assume that players do not possess all relevant information a priori but instead try to infer (learn) optimal behavior from past experience. Therefore players receive feedback about play at the end of each repetition. We maintain our previous assumption, i.e. private beliefs are not revealed and at the end of each round each player observes the entire sequence of actions and the state of Nature. However, we will now assume in addition that the private belief distribution chosen in a round is *not* part of any player’s feedback. We thus add another source of uncertainty regarding the social learning setting and we will frequently refer to such an environment as one with *fundamental structural uncertainty*. We further discuss this assumption below. Finally, as before players are assumed to keep track of frequencies  $\kappa(h_i, \theta \mid \zeta_r)$  across *all* repetitions, form assessments according to (2) for  $\epsilon \rightarrow 0$  and myopically Bayes rationally respond to these assessments as defined by (4).

#### Motivation and Interpretation

The assumptions above entail that players do not distinguish between social learning settings (i.e. private belief distributions) while adapting. More precisely neither is the relevant distribution in a round revealed to players, nor do players try to derive it from available information, nor do players form a subjective belief about it which they update over time. In the beginning of this subsection we have raised our main argument in favor of these assumptions: The prevalent concern that a single

---

<sup>29</sup>Hence, wlog  $\pi_k > 0$  for each  $k = 1, \dots, K$  and  $\sum_{k=1}^K \pi_k = 1$ .

setting does not offer sufficient opportunities for successful adaptation. Further arguments in favor of our assumptions may be raised.

Notice first that we have restricted ourselves to a set of social learning settings which differ only in the private belief distributions. More generally one could assume that settings differ in other fundamentals as well. We decided to focus on the private belief distribution since it seems the element which is hardest to identify.<sup>30</sup> Indeed we have serious doubts about players knowing the distribution from which other players draw their private beliefs or being able to pin down a set of likely distributions and form a prior about it.

Secondly, it is not clear how direct feedback about the distribution of private beliefs can be provided unless as argued before one is willing to assume that physical cues are correlated with the information structure. Players would therefore have to rely on indirect measures. Indeed, the best indicator of the realized distribution is provided by a player's own private belief.<sup>31</sup> It might thus be argued that in an extended global social learning game one should assume players to condition their assessments on their own private beliefs as well. However, this would require players to keep track of the complete sequence of (own) private beliefs across repetitions. This is for instance not possible if adaptation takes place in a population model with new players interacting each round if beliefs are not transmitted to subsequent generations.<sup>32</sup> Additionally players need to believe that own private beliefs constitute an important source of information when learning about the informativeness of others' decisions. *Naivety* in the spirit of Esponda (2008a) may preclude such beliefs and lead players to deliberately ignore own private beliefs when forming assessments.

## Analogy-Based Expectations Equilibrium

In order to discuss convergence of the process we now introduce the concept of analogy-based expectations equilibrium proposed in Jehiel (2005) for multi-stage games with (almost) perfect information and later on extended to static (Jehiel and Koessler, 2008) and dynamic (Jehiel and Ettinger, 2010) games of incomplete information.

We note first that in the global social learning game tree a *node* of player  $j$  is characterized by a tuple  $(k, \theta, s_1, \dots, s_i, h_i)$  where  $k$  is the social learning setting,  $\theta$  is the realized state of Nature,  $s_1, \dots, s_i$  is the sequence of private beliefs up to and including player  $i$  and  $h_i$  is the history of previous choices. Let  $X_j$  denote the set of nodes for player  $j$ . Representation of player  $j$ 's information sets  $(s_j, h_j)$  as collections of such nodes is straightforward. In general an *analogy partition*  $\mathcal{A}_i$  for player  $i \in \{1, \dots, n\}$  is a partition of the set  $\{(j, x_j) : j \neq i, x_j \in X_j\}$  into subsets  $\alpha_i$  called *analogy classes*.<sup>33</sup>  $(\mathcal{A}_i)_{i=1}^n$  denotes a profile of analogy partitions. We will consider a single specific such

---

<sup>30</sup>This does not reflect a conviction, that players will necessarily be able to identify the other properties of the game. Indeed, it is not clear, how a player might identify the correct environment when she is about to act.

<sup>31</sup>In some environments it may also be possible to identify the setting from differences in behavior. The arguments raised below extend to such possibilities as well.

<sup>32</sup>This is the learning framework considered in recurring games (Jackson and Kalai, 1997) and employed for instance by Jehiel and Koessler (2008) to motivate their equilibrium concept. Note however Miettinen (2010) arguing for the standard learning framework that own types may not be remembered over time in particular when these are not linked to personal characteristics.

<sup>33</sup>A partition of a set  $X$  is a collection of subsets  $Y_k \subseteq X$  such that  $\bigcup_k Y_k = X$  and  $Y_k \cap Y_{k'} = \emptyset$  for  $k \neq k'$ .

profile. In particular for each player  $i$   $\mathcal{A}_i = \{\alpha_i(\theta, h_j) : h_j \in \bigcup_{j < i} H_j, \theta \in \Theta\}$  where formally  $\alpha_i(\theta, h_j) = \{(k', \theta', s'_1, \dots, s'_j, h'_j) \in X_j : \theta' = \theta, h'_j = h_j\}$ . Hence, players differentiate others' nodes according to the state of Nature and the observed history of previous choices, but not according to the social learning setting or the player's private belief. We refer to this specific profile of partitions as the *information-anonymous analogy partition*.

An *analogy based expectation* for player  $i$  is given by a mapping  $\bar{\sigma}_i : \mathcal{A}_i \rightarrow \Delta(A)$  assigning to each analogy class a probability measure on action space  $A$ . In a slight abuse of notation we let  $\bar{\sigma}(h, \theta)$  denote the probability of investment expected in class  $\alpha_i(\theta, h_j)$ . Given that the social learning setting is not revealed to players behavior is captured as in the standard social learning game by *behavioral strategies*  $\sigma_i : \bigcup_{k=1}^K \text{supp}(G_1^k) \times H_i \rightarrow \Delta(A)$  where  $\sigma_i(s_i, h_i)$  denotes player  $i$ 's probability of investment at private belief  $s_i$  and history  $h_i$ . A profile of strategies is denoted by  $\sigma$ . In equilibrium players will be required to choose strategies which are (sequential) best responses to *beliefs* derived from analogy-based expectations in a Bayesian way. Accordingly, beliefs of player  $i$  are given by the mapping  $\bar{b}_i : \bigcup_{k=1}^K \text{supp}(G_1^k) \times H_i \rightarrow \Delta(\Theta)$  where  $\bar{b}_i(s_i, h_i)$  denotes the estimated probability that the true state of Nature is 1 given private belief  $s_i$  and history  $h_i$ .

**Definition 4.** A strategy profile  $\sigma$  is an *Analogy-based Expectations Equilibrium with information-anonymous analogy partition (ABEE)* if and only if there exist analogy-based expectations  $(\bar{\sigma}_i)_{i=1}^n$  and a system of beliefs  $(\bar{b}_i)_{i=1}^n$  satisfying

$$(i') \text{ **Bayes' rule:}** \quad \bar{b}_i(s_i, h_i) / [1 - \bar{b}_i(s_i, h_i)] = s_i \bar{\varphi}_i(h_i | 1) / (1 - s_i) \bar{\varphi}_i(h_i | 0)$$

for each  $i = 1, \dots, n$ , each  $s_i$  and each  $h_i$  satisfying  $\bar{\varphi}_i(h_i | \theta) > 0$  for some  $\theta \in \Theta$  where

$$\bar{\varphi}(h_1 | \theta) = 1 \quad \text{and} \quad \bar{\varphi}_i((h_j, a_j) | \theta) = \bar{\varphi}_i(h_j | \theta) * \bar{\sigma}_i(a_j | \theta, h_j)$$

for each  $\theta \in \Theta$  and each  $1 < j < i$ ,  $h_j \in H_j$  and  $a_j \in A$ ,

(ii') **Consistency of analogy-based expectations:** for each  $i = 1, \dots, n$ ,  $h_j \in H_j$  and  $\theta \in \Theta$

$$\bar{\sigma}_i(\theta, h_j) = \sum_{k=1}^K \int_{\text{supp}(G_\theta^k)} \nu_\sigma(k, ds_j | h_j, \theta) \sigma(s_j, h_j)$$

where  $\nu_\sigma$  is the distribution on tuples  $(k, s_j, h_j, \theta)$  induced by  $\sigma$  and the fundamentals,

$$(iii') \text{ **Sequential Best Response:}** \quad \sigma_i(s_i, h_i) = \begin{cases} 1 & \text{if } \bar{b}_i(s_i, h_i) > 1/2 \\ 0 & \text{if } \bar{b}_i(s_i, h_i) < 1/2 \end{cases}.$$

In the appendix we prove the following characterization of analogy-based expectations equilibria.

**Proposition 3.** For any ABEE  $\sigma^A$  it holds for each  $i = 1, \dots, n$ , each  $h_j \in H_j$  and each  $\theta \in \Theta$

$$\bar{\varphi}_i(h_j | \theta) = \bar{\varphi}(h_j | \theta) = \sum_{k=1}^K \pi_k \varphi^k(h_j | \theta)$$

where for each  $k = 1, \dots, K$  and each  $h_j, a_j$  and  $\theta$

$$\varphi^k(h_1 | \theta) = 1 \quad \text{and} \quad \varphi^k((h_j, a_j) | \theta) = \varphi^k(h_j | \theta) * \int_{s_j \in \text{supp}(G_\theta^k)} \sigma_j^A(s_j, h_j) dG_\theta^k(s_j).$$

The set of ABEE is the set of strategy profiles which are iteratively undominated with respect to the ex-ante expected utilities  $U_i(\sigma_i, \bar{\sigma})$  derived from analogy-based expectations. Furthermore besides in a non-generic class of global social learning games the ABEE outcome is unique.

Subsequently we refer to the probabilities  $\{\bar{\varphi}(h | \theta)\}_{h \in H, \theta \in \Theta}$  as *analogy-based assessments*. Furthermore we denote an ABEE strategy profile by  $\sigma^A$ . As before we admit to neglect those non-generic settings in which ties arise. Therefore probabilities  $\bar{\varphi}(h | \theta)$  and strategic behavior  $\sigma^A(s, h)$  at histories satisfying  $\bar{\varphi}(h | \theta) > 0$  for some (each)  $\theta \in \Theta$  are uniquely defined.

## Convergence

The standard interpretation of ABEE is as the limiting outcome of a learning process possibly involving imperfect feedback about others. Our next Proposition (proven in the appendix) corroborates this interpretation.

**Proposition 4.** *Take an extended global social learning game and let the learning horizon be infinite. Along almost any learning path  $\zeta_\infty$  of the extended adaptive process and for any  $\epsilon > 0$  eventually assessments are  $\epsilon$ -close to analogy-based assessments and strategic responses play  $\epsilon$ -like the ABEE strategy profile at all histories occurring with strictly positive probability.*

## ABEE and Rational Play

**Corollary.** *Generically for some  $k \in \{1, \dots, K\}$  the ABEE strategy differs from the rational strategy for the standard social learning game  $(n, A, u, \Theta, (G_0^k, G_1^k))$ .<sup>34</sup>*

The corollary ascertains that Bayesian rational players which adapt across social learning settings will generically not achieve maximal expected payoff in each setting. In the appendicized proof we provide a generic example for this assertion. Yet, Bayesian rational responses may still eventually lead to optimal behavior taking into account the restricted structural knowledge of players. Our main result below (proven in the appendix) invalidates this claim. Bayesian rational players who adapt across settings may arrive at a suboptimal strategy even in the global game.

**Lemma 2.** *Given the vector of assessments  $(\varphi^1, \dots, \varphi^K)$  the benchmark strategy  $\sigma^*$  given by*

$$\sigma_{\mathcal{E}}^*(s, h) = \begin{cases} 1 & \text{if } \sum_{k=1}^K \alpha_k \bar{\varphi}_k^*(h | 1) dG_1^K(s) > \sum_{k=1}^K \alpha_k \bar{\varphi}_k^*(h | 0) dG_0^K(s) \\ 0 & \text{if } \sum_{k=1}^K \alpha_k \bar{\varphi}_k^*(h | 1) dG_1^K(s) < \sum_{k=1}^K \alpha_k \bar{\varphi}_k^*(h | 0) dG_0^K(s) \end{cases}$$

*maximizes expected payoff  $U_i(\sigma | \varphi^1, \dots, \varphi^K)$  on  $\Sigma$  for each  $i = 1, \dots, n$ .*

**Proposition 5.** *Generically the ABEE strategy does not coincide with the benchmark strategy at all histories reached a non-vanishing fraction of the time.<sup>35</sup>*

---

<sup>34</sup>That is  $\sigma^A(s, h) \neq \sigma^{*k}(s, h)$  for some  $h \in H$  and each private belief  $s$  in some subset  $B \subseteq \text{supp}(G_\theta^k)$  where  $\int_B dG_\theta^k(s) > 0$  and  $\bar{\varphi}(h | \theta) + \varphi^{*k}(h | \theta) > 0$  for some  $\theta \in \Theta$ .

<sup>35</sup>That is  $\sigma^A(s, h) \neq \sigma^*(s, h)$  for some  $h \in H$  and each  $b(s, \emptyset)$  in some  $B \subseteq \bigcup_{k=1}^K [\underline{b}_k, \bar{b}_k]$  where  $\sum_{k=1}^K \pi_k \int_{s \in B} dG_\theta^k(s) > 0$  and  $\sum_{k=1}^K \pi_k \varphi^k(h | \theta) > 0$  for some  $\theta \in \Theta$ .

## ABEE and Non-Bayesian Strategies

Proposition 5 demonstrates how Bayesian rational players when forced to adapt across settings may be led to a suboptimal strategy. An implication of this result is that Bayesian rationality itself is subject to challenge in such environments. For instance players who experiment actively may discover non-Bayesian rational strategies which perform better. Alternatively evolutionary forces may select such strategies.<sup>36</sup> The nature of behavior which may outperform the Bayesian rational one is thus of interest and we aim at shedding more light on this issue.

For shortness of the exposition we relegate a detailed discussion to the appendix (Appendix C). Our analysis focuses on a restricted class of simple strategies related to an extended model of social learning suggested recently (March and Ziegelmeyer, 2009) with the aim to reconcile theoretical predictions and experimental evidence. In this model players differ with respect to the weight they place upon their private information relative to the information derived from observed actions of others when forming beliefs. Formally, a player with *private information weight*  $\beta$  forms beliefs according to

$$b(s, h, \beta) = \frac{[b(s, \emptyset)]^\beta b(\emptyset, h)}{[b(s, \emptyset)]^\beta b(\emptyset, h) + [1 - b(s, \emptyset)]^\beta [1 - b(\emptyset, h)]}. \quad (6)$$

Accordingly, if  $\beta = 1$  the player is Bayesian, while if  $\beta > 1$  ( $\beta < 1$ ) the player overweights (underweights) her private information. March and Ziegelmeyer (2009) show how the extended model is able to capture the experimental evidence if sufficiently many overweighers are present in the population. We consider a finite set of parameters  $\beta$  symmetric around  $\beta = 1$  and investigate numerically the limiting outcome of the discrete replicator dynamics for a population over this set. Our analysis therefore contributes to an explanation of the experimental results, a point which we return to later.

Our results are summarized in the following result. For a more detailed exposition of the results and a discussion of the underlying main assumptions we refer the interested reader to the appendix.

**Result 1.** *When adaptation takes place across games both moderate over- and underweighting robustly survive in the evolutionary process. In particular moderate overweighers of private information are more likely to be present in a population (i) the shorter the social learning sequence, (ii) the larger the precision of private signals, and (iii) the larger the difference of the basic social learning settings. On the other hand moderate underweighers are more likely to be present, the longer the sequence, the lower the signal precision, and the larger the difference between settings.*

## 6.2 Finite Learning Horizon

Despite individuals learning across a variety of games the *medium run* of a learning process may still be a more relevant predictor of behavior. First, many strategic environments do not remain unchanged in the very long term. Second, active learning is costly which means that players will likely stop experimenting and settle on a strategy after a finite number of repetitions. Third, the structure of social learning games implies that decision nodes late in the sequence will be encountered exponentially less frequently than nodes at the very beginning.<sup>37</sup> Hence, even if enough information has been accumulated in order to make optimal decision in early periods, players will still be weakly adapted

<sup>36</sup>That is provided expected payoffs and evolutionary fitness are sufficiently correlated.

<sup>37</sup>There exist  $2^i$  different history-state-pairs in period  $i$  of the social learning game.

in later periods. Finally, cognitive abilities are costly as well which means that individuals may be better characterized by finite memory and information processing capabilities (see for instance Jehiel, 1995; Rubinstein, 1986; Young, 1993). For instance Samuelson (2004, 2006) argues that endowing humans with better information processing capabilities is prohibitively costly. In this subsection we will therefore explicitly study adaptation with a finite horizon. Since we have addressed adaptation across games previously we focus on the case of a single, fixed social learning game.

We invoke the following main assumption: For any history  $h \in H_{-1} = H \setminus H_1$ <sup>38</sup> each player holds a noisy assessments of the public belief given by the distorted public belief  $b'(\varnothing, h) = b(\varnothing, h) + \tilde{\epsilon}_h$ . Distortions  $\tilde{\epsilon}_h$  are independent across histories  $h \in H_{-1}$  and normally distributed with mean zero and standard deviations  $(\eta_h)_{h \in H_{-1}}$  such that  $Pr(0 < b'(\varnothing, h) < 1) \approx 1$  for each  $h$ .

To defend this formalization notice that public beliefs satisfy  $b(\varnothing, h) = E[\tilde{\theta} | h]$ . This relation justifies a reinterpretation of the learning process. In particular players can learn public beliefs by calculating *sample means*

$$b'(\varnothing, h) = \bar{\theta}(h, \zeta_R) = \frac{1}{\kappa(h | \zeta_R)} \sum_{r \leq R; h \in y(r)} \theta(r)$$

where  $\kappa(h, \zeta_R)$  denotes the number of occurrences of history  $h$  in super-history  $\zeta_R$ . By the central limit theorem (see for instance Shao, 2003)  $\bar{\theta}(h, \zeta_R)$  is asymptotically normally distributed with mean  $b(\varnothing, h)$  and variance  $\eta_h^2 = b(\varnothing, h)(1 - b(\varnothing, h))/\kappa(h | \zeta_R)$ . Accordingly, with a finite number of repetitions assessments of public beliefs remain noisy ( $\eta_h^2 > 0$ ).

The existence of simple white noise in assessments remains our only assumption. Our first result proven in the appendix shows how this assumption affects Bayesian posterior beliefs.

**Proposition 6.**<sup>39</sup> *For any history  $h \in H_{-1}$  and private belief  $b(s, \varnothing)$ , if  $\eta_h > 0$ ,*

$$E_{\tilde{\epsilon}_h} [b'(s, h)] = \begin{cases} > b(s, h) & \text{if } b(s, \varnothing) < \frac{1}{2} \\ < b(s, h) & \text{if } b(s, \varnothing) > \frac{1}{2} \end{cases}$$

Due to the convexity (resp. concavity) of the Bayesian updating rule in the domain  $b(s, \varnothing) < (\text{resp. } >) \frac{1}{2}$ , contradictory information is weighed more heavily than confirmatory information. Since noisy public beliefs are equally likely to differ from correct ones in either direction, posteriors are biased in the direction of the public information. In other words players' posteriors (in expected terms) lean more towards the public information, i.e. players overweight private information.

In a world with rich action space where players are rewarded for stating posteriors, the Lemma above implies that noisy assessments favor overweighting of private information.<sup>40</sup> With coarse action set  $A = \{0, 1\}$ , this implication is not true. Our main result of this subsection more generally proves that Bayes-rationally responding to noisy assessments is not the optimal strategy.

In order to establish this result let  $\langle n, \Theta, A, u, (G_0, G_1), \boldsymbol{\eta} \rangle$  denote the *noisy social learning game* where  $\boldsymbol{\eta} = (\eta_h)_{h \in H}$  such that  $\eta_{h_1} = 0$  is the vector of standard deviations and we identify the normal distribution with mean and standard deviation zero with the (discrete) Dirac measure at zero. The noisy social learning game differs from the standard game by the restriction that players instead of

<sup>38</sup>Obviously, there is no room for noisy assessments in the first period.

<sup>39</sup>Appendix B derives a similar result for a setting with normally distributed state and private signals which is more closely related to the literature on learning in rational expectations.

<sup>40</sup>See Appendix B.



histories  $h \in H$  observe realizations of  $b'(\emptyset, h) = b(\emptyset, h) + \tilde{\epsilon}_h$ . Accordingly behavioral strategies of players are given by mappings  $\sigma_\eta : [\underline{b}, \bar{b}] \times [0, 1] \rightarrow [0, 1]$  where  $\sigma_\eta(s, x) = \Pr(a = 1 \mid b(s, \emptyset), b'(\emptyset, h) = x)$ . Hence, we presume that a player's public information is completely captured by her (distorted) public belief. In particular if a player's perceived public belief is the same at different histories in (possibly) different periods she has to employ the same strategy mapping private beliefs into actions. Taking into account that players are assigned randomly to periods a player's (ex-ante) expected payoff to strategy  $\sigma$  given probabilities  $\Pr(h \mid \theta)$ <sup>41</sup> is then given by

$$U_\eta(\sigma) = \frac{1}{4n} \int_{\underline{b}}^{\bar{b}} \sum_{h \in H} \int_{-b(\emptyset, h)}^{1-b(\emptyset, h)} \sigma_\eta(s, b(\emptyset, h) + \epsilon) [\Pr(h \mid 1) dG_1(s) - \Pr(h \mid 0) dG_0(s)] \phi_h(\epsilon) d\epsilon$$

where  $b(\emptyset, h) = \Pr(h \mid 1) / [\Pr(h \mid 1) + \Pr(h \mid 0)]$ , and  $\phi_h$  denotes the density of  $\mathcal{N}(0, \eta_h)$ <sup>42</sup>.

**Lemma 3.** *Given probabilities  $\Pr(h \mid \theta)$  for  $\theta = 0, 1$  let*

$$t(x \mid \eta) = \frac{\sum_{h \in H} \Pr(h \mid 0) \phi_h(x - b(\emptyset, h))}{\sum_{h \in H} [\Pr(h \mid 0) + \Pr(h \mid 1)] \phi_h(x - b(\emptyset, h))}.$$

*The benchmark strategy  $\sigma_\eta^*$  given by  $\sigma_\eta^*(s, x) = 1$  if  $s > t(x \mid \eta)$  and  $\sigma_\eta^*(s, x) = 0$  if  $s < t(x \mid \eta)$  maximizes expected payoff.*

Let Bayesian rational strategy be given by  $\sigma^*$  where  $\sigma^*(s, x) = 1$  if  $s > 1 - x$  and  $\sigma^*(s, x) = 0$  if  $s < 1 - x$ . As in the previous subsection we show that generically this strategy does not maximize expected payoff.

**Proposition 7.** *Generically for sufficiently large  $\eta$  the Bayes-rational strategy  $\sigma^*$  does not coincide with the benchmark strategy  $\sigma_\eta^*$  at all histories reached a non-vanishing fraction of the time.*<sup>43</sup>

Proposition 7 provides the counterpart to Proposition 5 in the previous subsection. A similar discussion regarding the nature of alternative behavior as the one following the result there is thus equally relevant at this point. Again we refer the reader to appendix C for a detailed analysis of this issue and restrict ourselves here to stating the main result.

**Result 2.** *If the learning horizon is finite such that players' assessments remain noisy, both over- and underweighting of private information robustly survive in an evolutionary process. In particular settings which involve a shorter learning horizon and a lower precision of private information are more favorable to overweighting while underweighting is favored by a longer learning horizon, higher precision of private information, and longer social learning sequences. In addition a finite learning horizon may favor extreme forms of overweighting.*

<sup>41</sup>Notice that these probabilities in general depend on the distribution of strategies in the population.

<sup>42</sup>Notice that  $\phi_{h_1}(0) = 1$  and  $\phi_{h_1}(x) = 0$  for  $x \neq 0$ .

<sup>43</sup>That is  $\sigma^*(s, x) \neq \sigma_\eta^*(s, x)$  for each  $x \in X \subseteq [0, 1]$  and each  $b(s, \emptyset) \in B \subseteq [\underline{b}, \bar{b}]$  where  $X$  and  $B$  satisfy  $\int_B dG_\theta(s) > 0$  and  $\int_X \sum_{h \in H} \Pr(h \mid \theta) \phi_h(x - b(\emptyset, h)) dx > 0$  for some  $\theta \in \Theta$ .

## 7 Discussion

### 7.1 Related Literature

The paper is related to several strands of literature. First, it extends the literature on social learning which developed mainly in the last two decades beginning with Bikhchandani, Hirshleifer, and Welch (1992) and Banerjee (1992). Those papers have given rise to a huge body of both theoretical and experimental literature. As discussed in Section 3 the majority of theoretical studies is characterized by a particularly strong view of Bayesian rationality.<sup>44</sup> On the other hand the experimental literature<sup>45</sup> establishes systematic deviations of subjects' behavior from rational play especially in situations where private and public information are conflicting. Indeed most subjects behave as if they overweight their private information.

Recently, the apparent gap between the findings of the theoretical and experimental literature on social learning has led to the development of alternative models which deviate from the prevalent Bayesian rational view<sup>46</sup>. Most of these models directly introduce biases into the decision process of players without providing a formal justification. They thus differ from the learning-oriented approach pursued in this paper.<sup>47</sup> In particular the focus of this paper is to investigate the learning foundations of economic models of social learning. The analysis of an alternative model of social learning which directly invokes the findings presented here can be found in a companion paper (March and Ziegelmeyer, 2009).

The paper also connects to the literature on learning in games (see Fudenberg and Levine, 1998, 2009, for an overview). In general learning models can be classified into three categories – rational, epistemic (or beliefs-based), and behavioral learning (see Walliser, 1998; Hart, 2005). Our adaptive process is strongly motivated by ideas developed in the fictitious play approach (Brown, 1951) and belongs to the class of epistemic learning models. Notice that unlike standard normal-form game applications of fictitious play we study adaptation in an extensive-form game with feedback constraints. Epistemic learning assumes a moderate degree of individuals' sophistication. Individuals are less sophisticated in behavioral learning models where they stochastically choose strategies according to their performance in the past. Conceptual differences aside it has been shown (Hopkins, 2002; Camerer and Ho, 1999) that epistemic and behavioral learning models are close from a mathematical point of

---

<sup>44</sup>Smith and Sørensen (2000) provides the most comprehensive and exhaustive analysis of rational social learning in situations where players observe the full sequence of past decisions and the timing of decisions is exogenous. More general settings are for instance studied in Chamley (2004a) and Acemoglu, Munther, Lobel, and Ozdaglar (2010). Chamley (2004b) and Vives (2008) provide comprehensive overviews.

<sup>45</sup>The experimental literature on social learning starts with the seminal paper by Anderson and Holt (1997). A collection of contributions to this literature can be found in the recent meta-study by Weizsäcker (2010).

<sup>46</sup>Bernardo and Welch (2001); Kariv (2005); Eyster and Rabin (2010); Guarino and Jehiel (2009); Bohren (2010).

<sup>47</sup>Exceptions are Guarino and Jehiel (2009) and Bernardo and Welch (2001). Guarino and Jehiel (2009) also invokes the concept of ABEE interpreted as arising from a learning process with limited feedback. Our paper differs in that we provide a formal proof for the convergence of an adaptive process to the equilibrium. Bernardo and Welch (2001) study the persistence of overweighers in a model of group selection for a standard social learning setting. Since overweighers' actions tend to convey valuable private information the presence of such types can increase the expected payoff to the group. Bernardo and Welch establish conditions under which these benefits outweigh the costs to overweighers from choosing a suboptimal strategy. In contrast we demonstrate that overweighting may be *individually optimal* as a response to limited learning opportunities.

view. Moreover given that we are interested in conditions which prevent players from approaching fully-rational play, choice of a more sophisticated epistemic learning model would not seem to work to our advantage.

On the other hand epistemic learning is still rather naive in the way conjectures are updated and responded to. In the direction of higher sophistication alternative models permit active experimentation with suboptimal strategies or have players update beliefs in a Bayesian way (rational learning). We do not think that our main result – the fact that the presence of both strategic and structural uncertainty may prevent players from learning to behave optimally – is sensitive to the choice of the adaptive process. For instance assume that players are endowed with a prior on the strategy space and a prior on the space of the unknown parameters of the model (the space of private belief distributions) and across repetitions update these priors in a Bayesian way in the light of the feedback received. Assume that a history becomes less informative. This change may either be attributed to a different distribution of strategies (more non-strategic players) or a different distribution of private beliefs (less informative). The feedback we have discussed does not permit players to distinguish between these effects. Therefore some uncertainty must prevail and by similar arguments as below this may lead to systematic and severe mistakes in the outcome of the learning process. While we leave a thorough investigation of this conjecture for future research we address the idea of experimentation in an extension of our learning process presented in Appendix A.

A standard assumption in learning models is the focus on a single, fixed game. Yet, as discussed before many authors argue that learning models should be understood as describing learning across similar games. In this paper we study explicitly learning across games. Few other studies address this issue. Steiner and Stewart (2008) study learning in a large class of simultaneous-move games with action sets fixed across games. Players receive feedback in the form of signals and players learn by extrapolating from similar past situations where similarity is measured by a fixed similarity function operating on payoffs. The authors show that extrapolation may lead to contagion of actions across games and unique long-run outcomes. Mengel (2009, see also Grimm and Mengel, 2009) studies reinforcement<sup>48</sup> learning in a multiple strategic-form games environment with fixed action sets where players not only learn which actions to choose but also how to partition the set of games. Distinguishing games is costly, i.e. players incur a cost from partitioning the set of games which is larger the finer the partition. Mengel shows that generically players do not distinguish all games which in turn may (de)stabilize Nash equilibria which are (un)stable to learning in a single game. In addition learning across games may help to explain certain experimental phenomena. The difference between the two papers and ours is that we study an extensive-form game of incomplete information and define games to be similar if they differ only in the information structure. Moreover we justify learning across games by players' limited feedback which does not permit them to easily distinguish games.

In order to describe the long-run outcome of our adaptive process when adaptation takes place across games, we have used the concept of analogy-based expectations equilibrium (Jehiel, 2005; Jehiel and Koessler, 2008; Jehiel and Ettinger, 2010). This has a close relationship to other equilibrium concepts like self-confirming equilibrium (Fudenberg and Levine, 1993a; Dekel, Fudenberg, and Levine, 2004) or conjectural equilibrium (Battigalli, 1997, see also Kalai and Lehrer, 1993). A common feature these concepts share is that they are commonly interpreted as the limit outcome of a learning process

---

<sup>48</sup>It is shown that the results hold as well for stochastic fictitious play.

where players are subject to feedback constraints. The main difference of our paper is that we explicitly provide such a learning process and prove its convergence.<sup>49</sup>

## 7.2 Feedback and Player Heterogeneity

We have assumed so far that at the end of each round every player's feedback is given by the complete sequence of choices and the realized state of Nature. Therefore a player's feedback is independent of both the position the player occupies and the player's own choice. Both these assumptions can be relaxed.

Formally, let  $y_i(r) = y_i(h_{n+1}(r), \theta(r), s_1(r), \dots, s_n(r))$  denote the feedback received by the player acting at position  $i$  in round  $r$  where  $y_i(r)$  is a deterministic function of the sequence of choices, the state of Nature and the sequence of realized private beliefs. Unlike before we now allow for different feedback depending on the position a player occupies.<sup>50</sup> Fix player  $i$  and write  $x \in y_i(r)$  if  $x$  is comprised in  $y_i(r)$ . We distinguish two cases: First, while clearly  $h_{i+1}(r) \in y_i(r)$  it is not clear whether a player will observe at the end of the round actions chosen after her own choice as well. We therefore make the distinction between *complete feedback* if  $h_{n+1}(r) \in y_i(r)$  and *partial feedback* if  $(a_{i+1}(r), \dots, a_n(r)) \notin y_i(r)$ . Second, since a player always recalls her own choice and payoffs are realized at the end of each round it cannot be ruled out that a player knows the state of Nature at the end of a round (since at least some action's payoff must strictly depend on the realized state of Nature). Yet, given the specific form of the payoff function this is not necessarily the case in every round. More precisely, if the player rejects absent further information she does not know the true payoff investing would have yielded. We therefore distinguish *unconditional feedback* if  $\theta(r) \in y_i(r)$  for any  $i$  from *conditional feedback* given by  $\theta(r) \in y_i(r)$  iff  $a_i(r) = 1$ .<sup>51</sup>

Another restrictive assumption we have made concerns the Nature of private beliefs. In particular we have assumed that in each round (respectively each round the same game is played) every player's private belief is drawn from a single fixed private belief distribution. Independent repeated draws from a single distribution effectively average out all differences in players' information. But heterogeneity of players' information is a reasonable assumption since a player's quality of information is likely to be endogeneously determined e.g. via her cognitive or material resources. More importantly for the questions addressed in this paper a player's information quality may also decisively influence her learning opportunities. For instance a player with constantly very informative private beliefs will create more deviations from herds than a player with constantly less informative beliefs.

In order to take this into account we propose the following extension on the distribution of private

---

<sup>49</sup>Indeed, convergence of a learning process in games other than complete information normal-form games is rarely investigated. Notable exceptions are Fudenberg and Levine (1993b) for Bayesian learning and Beggs (1993) for a variant of behavioral learning in Bayesian games with binary actions in which players adjust their cutoff-strategies over time.

<sup>50</sup>Notice that this assumption is in line with a model of learning within a large population of players governed by a random matching mechanism which we will frequently refer to. For an explanation and a discussion of alternative modeling assumptions see e.g. Fudenberg and Levine (1998, section 1.2).

<sup>51</sup>Our distinctions are comparable to Esponda's (2008b) in the context of auctions. Obviously further restrictions are possible. For instance if players observe only aggregated frequencies of actions together with the realized state of Nature, the learning process may be expected to approach the coarse analogy-based expectations equilibrium (Guarino and Jehiel, 2009). Yet, since our main focus is a different one, we leave a complete characterization of the outcome of the learning process in dependence upon players' feedback for future research.

beliefs applying the *signal quality structure* model of Smith and Sørensen (2008b, section 5.2): Each player is told one of two signals  $\tilde{s} \in \{0, 1\}$  where  $Pr(\tilde{s} = \theta \mid \tilde{\theta} = \theta) = \tilde{q}$  denotes the signal precision. We let  $\tilde{q}$  be a random variable, distributed on the set  $(\frac{1}{2}, 1)$  according to the measure  $Q$ . For instance the Dirac measure on some  $q \in (0.5, 1)$  yields the specific model of Bikhchandani, Hirshleifer, and Welch (1992). We now assume that each individual is characterized by her measure  $Q$  which is private information and fixed across rounds. Accordingly, the population is characterized by a (finite discrete) probability distribution  $\mathcal{Q}$  on the set  $\Delta((1/2, 1))$  of all measures  $Q$  satisfying

$$\int_{Q \in \Delta((1/2, 1))} \int_{q=1/2}^1 \sum_{s \in \{0, 1\}} Pr(s \mid \theta, q) \mathbb{1}_A \left( \frac{Pr(s \mid 1, q)}{Pr(s \mid 1, q) + Pr(s \mid 0, q)} \right) Q(dq) \mathcal{Q}(dQ) = \int_{s \in A} dG_\theta(s)$$

for each  $A \subseteq [\underline{b}, \bar{b}]$  with  $\mathbb{1}_A$  the indicator function of the set  $A$ . We distinguish a *homogeneous support of private beliefs* if  $\mathcal{Q}$  is a Dirac measure on some  $Q \in \Delta((0.5, 1))$  from a *heterogeneous support of private beliefs* if it is not.

**Proposition 8.** *Propositions 2 and 4 continue to hold with partial and conditional feedback and heterogeneous support of private beliefs.*

*Proof.* Both propositions assume an infinite learning horizon. A player will thus play (in every game) in position  $n$  an infinite number of times. Moreover even with the least informative support of private beliefs each round each player is endowed with a private belief  $b(s, h) > \frac{1}{2}$  favoring investment with strictly positive probability. Therefore a finite sequence of players are all endowed with such beliefs with strictly positive probability each round as well. Hence, independent of her private belief support each player will infinitely often invest at period  $n$  in which case she observes both, the state of Nature and the complete sequence of choices and neither feedback constraints nor heterogeneity of private beliefs matter.  $\square$

As will be shown below feedback constraints and heterogeneous support of private beliefs become crucial in a setting with endogenous timing.

### 7.3 Preferences

We have so far restricted ourselves to a simple social learning game. In more general settings it may be argued that the adaptive process is even less likely to resolve all uncertainty completely. This in turn questions even more profoundly the optimality of Bayesian rational responses.

Consider for instance a world with heterogeneity in preferences as introduced by Smith and Sørensen (2000). They distinguish *rational* types whose payoff decisively depends upon the realized state of Nature from *crazy* types who always choose the same action and thereby introduce noise into the social learning process. Additionally, Smith and Sørensen (2000) assume common knowledge of the distribution of preference types. As before from an individual learning perspective this assumption is disputable. In particular it would require players to receive feedback about payoffs of other players. Without this information identification of single environments is not possible and adaptation takes place across settings. Therefore presence of multiple social learning settings which differ in their distribution of preferences may trivially explain deviations from rational behavior in single settings. However, it can be shown that absent differences in private belief distributions the Bayes-rational strategy remains optimal if no player can distinguish settings.

Yet, differences in preferences may reinforce suboptimality of the Bayes-rational strategy in the presence of multiple private belief distributions. To see this consider a world composed of two simple social learning games occurring with equal probability. In each game,  $k \in \{1, 2\}$ ,  $\Theta = A = \{0, 1\}$ . Furthermore in each game every player may receive one of two possible private beliefs  $s_i \in \{1 - q_k, q_k\}$  where  $Pr(\tilde{s}_i = 1 - q_k \mid \tilde{\theta} = 0) = Pr(\tilde{s}_i = q_k \mid \tilde{\theta} = 1) = q_k$  and  $1 > q_k > 1/2$ . We assume that  $q_2 > q_1$ . In addition the games differ in their utility function. In particular in game 1 all players have standard preferences maintained so far (i.e.  $u(1, \theta) = \theta - 1/2$  and  $u(0, \theta) = 0$ ) while in game 2 only a fraction  $(1 - 2\xi_2)$  of players have these preferences while a fraction of each  $\xi_2 > 0$  are noise types which always invest or always reject respectively. In particular assume that  $\xi_2$  is sufficiently large such that players should optimally follow private information even after observing the first two players making the same choice. If players learn across these environments absence of noise in game 1 leads Bayes-rational players in game 2 to infer too much from the first players' decisions and may cause them to suboptimally imitate in period 3. On the contrary it can be shown that there robustly exist settings where overweighters of private information outperform Bayesian rational players by avoiding this mistake (see appendix D).

While noise players are an extreme form of preference heterogeneity similar arguments apply to the case of several rational types as has been most impressively demonstrated by the confounding outcomes of Smith and Sørensen (2000). In general heterogeneity in preferences complicates learning from others' actions by significantly reducing the amount of information single decisions convey. This brings into focus the problem of imperfect learning within a finite number of repetitions much more than in the simple settings studied so far. If heterogeneity confounds the opportunities to learn by observing others and imperfect learning opportunities add another source of noise, the process of social learning may break down altogether. Moreover as discussed above the inability of players to distinguish social learning environments may spread such complications across settings. Even if a single environment gives best chances to learning from others, these may be thwarted by player' adaptation across settings and their limited time to do so.

In conclusion imperfect learning opportunities may have much more drastic influence in settings where players must learn from subtle differences since these are most vulnerable to the introduction of any kind of noise. Accordingly such settings may be particularly favorable to non-Bayesian strategies.

## 7.4 Endogeneous Timing of Decisions

A different complication is introduced if players need to choose the timing of their decisions. In this case the very simple structure of the game which for instance drives our convergence result (Proposition 2) is lost. More precisely endogeneous timing of decisions requires players to not only derive information from past choices but also foresee future social learning opportunities. By choosing the timing of decisions players may then significantly influence their possibilities for individual learning. Moreover the importance of further restrictions of learning opportunities such as heterogeneous support of private beliefs, partial feedback on choices or conditional feedback on the state of Nature increases.

Consider for instance a setting where the costs of investing are large such that only sufficiently strong favorable information may induce a player to invest (e.g.  $u(1, \theta) = \theta - 2/3$ ). Assume further conditional feedback on the state of Nature, i.e. a player receives feedback about the profitability of the investment (the state of Nature) if and only if she invests herself. Suppose finally that some players are restricted to a limited quality of information such that neither of them will ever invest

conditional on private information alone. In the simple adaptive process this group of players will not be able to learn how to infer information from others' choices. Consequently neither of them will ever invest. Overweighting of private information may then be profitable not only because sometimes investing must yield a positive payoff but also because it would allow players to learn how to take others' actions into account in future interactions.

While this constitutes a rather extreme example, it facilitates how an endogenous timing of decisions complicates the task of learning from others and players' opportunities to adapt to this task. At least it shows how individual experimentation and self-confirming equilibria are much more important than in the present context.

## 7.5 Rule Rationality and Experimental Findings

According to our main result when players adapt across games they may develop alternative strategies which apply well to the multiple games environment. This idea has become prominent by the term *rule rationality* (Aumann, 1997, 2008, the opposite notion is *act-rationality*). In single decision tasks rule rational choices may lead to severe and systematic errors as demonstrated for instance in the corollary of section 6.1. Rule rationality is thus frequently employed to explain the divergence of theoretical predictions and experimental observations in both individual and strategic choice contexts.<sup>52</sup> Most commonly, rules are identified with simple heuristics (or "rules of thumb") which arise in and are subject to selection through learning, imitation, or evolution. In this spirit our results in appendix C demonstrate how non-Bayesian rules which for instance comprise the overweighting of private information may arise in a social learning context. As discussed above a stable finding in the experimental literature on social learning is that many players indeed behave as if they overweight their private information. Our results thus provide a new interpretation of these results. Subjects' behavior may deviate from rational behavior in the laboratory because these subjects employ rules well adapted to field environments where large uncertainty and limited learning opportunities persist.

## 8 Conclusion

This paper is a first attempt to discuss the learning foundations of rational play in social learning games. Our results suggest that too much structural knowledge has been assumed in standard economic models of social learning. Though in the absence of fundamental structural uncertainty and with an infinite learning horizon epistemic learning leads to Bayesian rational play, the same learning process favors non-Bayesian play whenever players do not know the distribution of private beliefs. As a consequence, further economic models of social learning should allow for the presence of fundamental structural uncertainty and learning models should inform those economic models about the nature and scope of the structural uncertainty.

---

<sup>52</sup>Proponents of this idea are e.g. Tversky and Kahnemann (1974), Myerson (1991), and more recently Al-Najjar and Weinstein (2009). Consider also the work on robustness to model misspecification in macroeconomic risk models. In these models having decision-makers take into account possible misspecifications of their model of the world induces model uncertainty premia on equilibrium prices of risk. Hansen (2007) for instance provides a nice overview of these results and discusses furthermore how model uncertainty can arise as the outcome of imperfect learning about a complex statistical problem.

Assuming that fundamental structural uncertainty is unavoidable, our theoretical study also suggests that *rule rationality* rather than *act rationality* is the appropriate benchmark for discussing rational social learning. If players lack structural knowledge then they are likely to develop rational *rules* of social learning rather than trying to act in a Bayesian rational way in each single social learning environment. This, in turn, seriously questions the informativeness of the existing experimental evidence on rational social learning. Indeed, the validity of the rational view of social learning should be tested in contexts that are familiar to economic actors, i.e. contexts with fundamental structural uncertainty. As far as we know, all existing experimental studies on social learning have considered laboratory settings which correlate strongly with the standard economic models of social learning. Though we understand that such simple settings with full control of information flows were the natural candidates for testing the existing economic models, we fear that the field environments in which social learning mainly takes place differ substantially from those laboratory settings. As a consequence, existing laboratory settings might be perceived as artificial by subjects and the fact that laboratory behavior systematically deviates from rational play in those settings might not come as a surprise and does not constitute conclusive evidence against rational social learning.

Our results therefore suggest a re-evaluation of both the experimental and the theoretical research in social learning. First, new laboratory experiments should be designed to test the rational view of social learning in more familiar contexts. Aspects of the field environment which are likely to strongly influence social learning (e.g. structural uncertainty) have to be incorporated in those laboratory settings. Second, theoretical models with players lacking structural knowledge are likely to provide a better understanding of real-world social learning. Notice that allowing for structural uncertainty in economic models of social learning does not necessarily increase the complexity of these models as shown for instance by March and Ziegelmeyer (2009).

Viewed from a broader perspective, our results also offer new insights for behavioral economists.<sup>53</sup> Despite a regular exchange between experimentalists and theorists over the past two decades, there is no satisfactory behavioral model of social learning. Previous attempts have imported psychological insights (e.g. judgmental biases, limited depth of reasoning) into existing economic models of social learning. These behavioral models acknowledge the cognitive limitations of economic actors by relaxing the assumption of Bayesian rationality in the direction of greater psychological realism. Though we are sympathetic to this approach, we show that a thorough investigation of the modeling assumptions may straightforwardly yield an alternative model of identical complexity but with increased explanatory power. We believe that our approach is likely to be fruitful not only in social learning. Economic models which accommodate the fact that field environments provide limited learning opportunities to players are likely to generate more accurate predictions without diminished tractability and in this sense they complement other models which incorporate more realistic psychological foundations.

---

<sup>53</sup>In recent years the field of economics has witnessed an increased emergence of behavioral models, many of them designed to explain experimental phenomena not captured by standard notions of equilibrium or rationality (see e.g. *Advances in Behavioral Economics* Camerer, Loewenstein, and Rabin, 2004). While these models usually better accommodate the evidence, a common critique concerns their lack of foundations (see e.g. Fudenberg, 2006).



## References

- ACEMOGLU, D., D. MUNTHE, I. LOBEL, AND A. OZDAGLAR (2010): “Bayesian Learning in Social Networks,” mimeo, MIT.
- AL-NAJJAR, N. I. (2009): “Decision Makers as Statisticians: Diversity, Ambiguity, and Learning,” *Econometrica*, 77, 1339–1369.
- AL-NAJJAR, N. I., AND J. WEINSTEIN (2009): “The Ambiguity Aversion Literature: A Critical Assessment,” *Economics and Philosophy*, 25, 249–284.
- ANDERSON, L., AND C. HOLT (1997): “Information cascades in the laboratory,” *American Economic Review*, 87, 847–862.
- AUMANN, R. (1997): “Rationality and Bounded Rationality,” *Games and Economic Behavior*, 21, 2–14.
- AUMANN, R. J. (2008): “Rule-Rationality versus Act-Rationality,” mimeo, The Hebrew University of Jerusalem.
- BANERJEE, A. V. (1992): “A simple model of herd behavior,” *The Quarterly Journal of Economics*, 107, 797–817.
- BATTIGALLI, P. (1997): “Comportamento Razionale ed Equilibrio nei Giochi e nelle Situazioni Sociali,” Ph.D. thesis, Bocconi University.
- BEGGS, A. (1993): “Learning in Bayesian Games with Binary Actions,” *The B.E. Journal of Theoretical Economics*, 9: Iss. 1 (Advances), Article 33.
- BERNARDO, A., AND I. WELCH (2001): “On the Evolution of Overconfidence and Entrepreneurs,” *Journal of Economics & Management Strategy*, 10, 301–330.
- BERNHEIM, B. D. (1984): “Rationalizable Strategic Behavior,” *Econometrica*, 52, 1007–1028.
- BIKHCHANDANI, S., D. HIRSHLEIFER, AND I. WELCH (1992): “A theory of fads, fashion, custom, and cultural change as informational cascades,” *Journal of Political Economy*, 100, 992–1026.
- (1998): “Learning from the behavior of others: Conformity, fads, and informational cascades,” *Journal of Economic Perspectives*, 12, 151–170.
- BOHREN, J. A. (2010): “Information-Processing Uncertainty in Social Learning,” Working paper, University of California, San Diego.
- BRANDENBURGER, A. M. (1996): “Strategic and Structural Uncertainty in Games,” in *Wise Choices*, ed. by R. Zeckhauser, Keeney, and Sebenius. Springer, Berlin.
- BROWN, G. W. (1951): “Iterative Solution of Games by Fictitious Play,” in *Activity Analysis of Production and Allocation*, ed. by T. Koopmans, pp. 374–376. Wiley, New York.
- CAMERER, C. F., AND T.-H. HO (1999): “Experience-weighted attraction learning in normal-form games,” *Econometrica*, 67, 827–874.

- CAMERER, C. F., G. LOEWENSTEIN, AND M. RABIN (eds.) (2004): *Advances in Behavioral Economics*. Russell Sage Foundation and Princeton University Press.
- CHAMLEY, C. (2004a): “Delays and Equilibria with Large and Small Information in Social Learning,” *European Economic Review*, 48, 477–501.
- (2004b): *Rational herds: Economic models of social learning*. Cambridge, Cambridge University Press.
- COCHRAN, W. G. (1934): “The distribution of quadratic forms in a normal system, with applications to the analysis of covariance,” *Math Proc of the Cambridge Philosophical Society*, 30, 178–191.
- DEKEL, E., D. FUDENBERG, AND D. K. LEVINE (2004): “Learning to Play Bayesian Games,” *Games and Economic Behavior*, 46, 282–303.
- DEKEL, E., AND F. GUL (1997): “Rationality and Knowledge in Game Theory,” in *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress*, ed. by D. M. Kreps, and K. F. Wallis, vol. 1 of *Econometric Society Monographs*, Vol. 26, pp. 87–172. Cambridge: Cambridge University Press.
- ESPONDA, I. (2008a): “Behavioral Equilibrium in Economies with Adverse Selection,” *The American Economic Review*, 98, 1269–1291.
- (2008b): “Information Feedback in First Price Auctions,” *RAND Journal of Economics*, 39, 491–508.
- EYSTER, E., AND M. RABIN (2010): “Naive Herding in Rich-Information Settings,” *American Economic Journal: Microeconomics*, forthcoming.
- FUDENBERG, D. (2006): “Advancing Beyond “Advances in Behavioral Economics”,” *Journal of Economic Literature*, 44, 694–711.
- FUDENBERG, D., AND D. M. KREPS (1993): “Learning Mixed Equilibria,” *Games and Economic Behavior*, 5, 320–367.
- (1995): “Learning in Extensive-Form Games, I: Self-Confirming Equilibria,” *Games and Economic Behavior*, 8, 20–55.
- FUDENBERG, D., AND D. K. LEVINE (1993a): “Self-Confirming Equilibrium,” *Econometrica*, 61, 523–545.
- (1993b): “Steady State Learning and Nash Equilibrium,” *Econometrica*, 61, 547–573.
- (1998): *The Theory of Learning in Games*. Cambridge, MA: MIT Press.
- (2009): “Learning and Equilibrium,” *Annual Review of Economics*, 1, 385–419.
- GALE, D. (1996): “What have we learned from social learning?,” *European Economic Review*, 40, 617–628.
- GRIMM, V., AND F. MENGEL (2009): “An Experiment on Learning in a Multiple Games Environment,” mimeo, Maastricht University.

- GUARINO, A., AND P. JEHIEL (2009): “Social Learning with Coarse Inference,” Working paper.
- GUESNERIE, R. (1992): “An Exploration of the Eductive Justifications of the Rational-Expectations Hypothesis,” *The American Economic Review*, 82, 1254–1278.
- HANSEN, L. P. (2007): “Beliefs, Doubts and Learning: Valuing Macroeconomic Risk,” *The American Economic Review*, 97, 1–30.
- HART, S. (2005): “Adaptive Heuristics,” *Econometrica*, 73, 1401–1430.
- HENDRICKS, K., AND D. KOVENOCK (1989): “Asymmetric Information, Information Externalities, and Efficiency: The Case of Oil Exploration,” *The RAND Journal of Economics*, 20, 164–182.
- HOPKINS, E. (2002): “Two Competing Models of How People Learn in Games,” *Econometrica*, 70, 2141–2166.
- JACKSON, M. O., AND E. KALAI (1997): “Social Learning in Recurring Games,” *Games and Economic Behavior*, 21, 102–134.
- JEHIEL, P. (1995): “Limited Horizon Forecast in Repeated Alternate Games,” *Journal of Economic Theory*, 67, 497–519.
- (2005): “Analogy-Based Expectation Equilibrium,” *Journal of Economic Theory*, 123, 81–104.
- JEHIEL, P., AND D. ETTINGER (2010): “A Theory of Deception,” *American Economic Journal: Microeconomics*, 2, 1–20.
- JEHIEL, P., AND F. KOESSLER (2008): “Revisiting Games of Incomplete Information with Analogy-Based Expectations,” *Games and Economic Behavior*, 62, 533–557.
- KALAI, E., AND E. LEHRER (1993): “Rational Learning Leads to Nash Equilibrium,” *Econometrica*, 61, 1019–1046.
- KARIV, S. (2005): “Overconfidence and Informational Cascades,” mimeo, UC Berkeley.
- LEE, I. H. (1993): “On the Convergence of Informational Cascades,” *Journal of Economic Theory*, 61, 395–411.
- MARCH, C. (2010): “A New Perspective on Social Learning,” Unpublished Ph.D. Thesis, Technical University of Berlin.
- MARCH, C., AND A. ZIEGELMEYER (2009): “Behavioral Social Learning,” mimeo, MPI of Economics, Jena.
- MENGEL, F. (2009): “Learning across Games,” mimeo, Maastricht University.
- MIETTINEN, T. (2010): “Paying Attention to Payoffs in Analogy-Based Learning,” *Economic Theory*, forthcoming.
- MILGROM, P., AND J. ROBERTS (1991): “Adaptive and Sophisticated Learning in Normal Form Games,” *Games and Economic Behavior*, 3, 82–100.

- MILGROM, P. R. (1981): “Good News and Bad News: Representation Theorems and Applications,” *Bell Journal of Economics*, 12, 380–391.
- MYERSON, R. (1991): *Game Theory: Analysis of Conflict*. Harvard University Press.
- PEARCE, D. G. (1984): “Rationalizable Strategic Behavior and the Problem of Perfection,” *Econometrica*, 52, 1029–1050.
- PEREA, A. (2001): *Rationality in Extensive Form Games*. Kluwer Academic Publishers, Boston.
- RUBINSTEIN, A. (1986): “Finite Automata Play the Repeated Prisoner’s Dilemma,” *Journal of Economic Theory*, 39, 83–96.
- SAMUELSON, L. (2004): “Information-Based Relative Consumption Effects,” *Econometrica*, 72, 93–118.
- (2006): “Information, Evolution and Utility,” *Theoretical Economics*, 1, 119–142.
- SHAO, J. (2003): *Mathematical Statistics*. Springer, New York, 2nd edn.
- SMITH, L., AND P. SØRENSEN (2000): “Pathological Outcomes of Observational Learning,” *Econometrica*, 68(2), 371–398.
- (2008b): “Rational Social Learning with Random Sampling,” Working paper.
- STEINER, J., AND C. STEWART (2008): “Contagion through Learning,” *Theoretical Economics*, 3, 431–458.
- TAN, T., AND S. WERLANG (1988): “The Bayesian Foundations of Solution Concepts of Games,” *Journal of Economic Theory*, 45, 370–391.
- TVERSKY, A., AND D. KAHNEMANN (1974): “Judgement under Uncertainty: Heuristics and Biases,” *Science*, 185, 1124–1131.
- VIVES, X. (1993): “How fast do rational agents learn?,” *Review of Economic Studies*, 60, 329–347.
- (1996): “Social Learning and Rational Expectations,” *European Economic Review*, 40, 589–601.
- (2008): *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton University Press.
- WALK, H. (2008): “A universal strong law of large numbers for conditional expectations via nearest neighbors,” *Journal of Multivariate Analysis*, 99, 1035–1050.
- WALLISER, B. (1998): “A Spectrum of Equilibration Processes in Game Theory,” *Journal of Evolutionary Economics*, 8, 67–87.
- WEIZSÄCKER, G. (2010): “Do we follow others when we should? A simple test of rational expectations,” *forthcoming in American Economic Review*.
- YOUNG, H. P. (1993): “The Evolution of Conventions,” *Econometrica*, 61, 57–84.
- ZIEGELMEYER, A., F. KOESSLER, J. BRACHT, AND E. WINTER (2010): “Fragility of information cascades: An experimental study using elicited beliefs,” *Experimental Economics*, 13, 121–145.

# **Adaptive Social Learning**

– Supplementary Material –

## A More Sophisticated Learning

In this addendum we extend the learning process to allow for players who are more sophisticated. We follow a general approach suggested by Fudenberg and Kreps (1995, FK henceforth). Since most of the concepts apply primarily to the long run we assume an infinite number of repetitions throughout.

For the adaptive process introduced in the main part of the text we have assumed that each player holds a single assessment  $\hat{\varphi}(\zeta_r)$  which she updates over time given evidence  $\zeta_r$ . More generally one may assume that a player considers possible several models of others. Hence, let  $\Phi = \{\varphi : H \times \Theta \rightarrow [0, 1]\}$  denote the set of possible assessments. Define a *general assessment rule*<sup>54</sup> as  $\hat{\gamma} : \bigcup_{r=1}^{\infty} \mathcal{Z}_r \rightarrow \Delta(\Phi)$  attaching to each learning path  $\zeta_r$  a probability distribution on  $\Phi$ . Within the general approach the updating of these general assessments over time is not modelled explicitly. This allows it to encompass various more specific models. Most importantly as shown by FK (Section 3.2) it encompasses Bayesian inference. Instead of defining explicit updating rules the general approach places conditions upon the relationship between assessments and evidence in the long run. For this purpose define for each  $r = 1, 2, \dots$ , each  $\zeta_r \in \mathcal{Z}_r$  and each  $h$  s.t.  $\kappa(h, \theta | \zeta_r) > 0$

$$\bar{\varphi}(h | \theta; \zeta_r) = \frac{\kappa(a, h, \theta | \zeta_r)}{\kappa(a, \theta | \zeta_r)}.$$

**Definition 5.** The general assessment rule  $\hat{\gamma}$  is **asymptotically empirical** if for every  $\epsilon > 0$ , every  $\zeta$  and for every  $\theta \in \Theta$  and  $h \in H$  such that  $\liminf_{r \rightarrow \infty} \frac{\kappa(h, \theta | \zeta_r)}{r} > 0$ ,

$$\lim_{r \rightarrow \infty} \hat{\gamma}(\zeta_r) (\{\varphi : \|\varphi(h | \theta) - \bar{\varphi}(h | \theta; \zeta_r)\| < \epsilon\}) = 1$$

A second cornerstone of our adaptive process are players' myopic responses to assessments. Apart from repeated play considerations this is also restrictive because it rules out the possibility that players experiment. For instance one can argue (see FK for an exhaustive discussion of this issue) that in extensive-form games a player's action affects what she learns about others' behavior. Hence, a more general approach should allow for experimentation of players. We take this into account as follows. For general assessment  $\gamma$  and some  $i \in \{1, \dots, n\}$ , let  $U_i(\sigma | \gamma) = \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \sigma(s, h_i) U(s, h_i | \gamma)$  where  $U(s, h | \gamma) = \int_{\varphi \in \Phi} [\varphi(h | 1) dG_1(s) - \varphi(h | 0) dG_0(s)] \gamma(d\varphi)$ . The strategic response  $\hat{\sigma}$  is **asymptotically myopic** with regard to  $\hat{\gamma}$  if there exists a sequence of non-negative numbers  $\{\epsilon_r\}_{r=1}^{\infty}$  such that  $\lim_{r \rightarrow \infty} \epsilon_r = 0$  and, for each  $r$  and  $\zeta_r$ , and each  $i = 1, \dots, n$

$$U_i(\hat{\sigma}(\zeta_r) | \hat{\gamma}(\zeta_r)) + \epsilon_r \geq \max_{\sigma \in \Sigma} U_i(\sigma, \hat{\gamma}(\zeta_r)).$$

Asymptotic myopia permits players to choose suboptimal strategies where the suboptimality vanishes over time. In particular players may choose slightly suboptimal strategies with larger probabilities or grossly suboptimal strategies with small probabilities (see (Fudenberg and Kreps, 1993)). Thus while at early dates depending on the sequence  $\{\epsilon_r\}_{r=1}^{\infty}$  the player may consciously experiment with suboptimal strategies, she eventually has to confine herself to *random experimentation* with decreasing overall probability. In this regard asymptotic myopia it still restrictive. A more general notion is the following

---

<sup>54</sup>This is not to be confused with the "rules" we discuss in connection with rule rationality.

**Definition 6.** The strategic response  $\hat{\sigma}$  is **asymptotically myopic with calendar-time limitations on experimentation** with respect to  $\hat{\gamma}$  if there exist

(i) a sequence  $\{\epsilon_r\}_{r=1}^{\infty}$  s.t.  $\epsilon_r > 0$  for each  $r = 1, 2, \dots$  and  $\lim_{r \rightarrow \infty} \epsilon_r = 0$ ,

(ii) a sequence  $\{\delta_r\}_{r=1}^{\infty}$  s.t.  $\delta_r \geq 0$  and  $\delta_{r+1} \geq \delta_r$  for each  $r = 1, 2, \dots$  and  $\lim_{r \rightarrow \infty} \delta_r/r = 0$ ,

(iii) an asymptotically myopic strategic response  $\hat{\sigma}^{opt}$ ,

(iv) a strategic response  $\hat{\sigma}^{exp}$  s.t. for every  $\zeta$ ,  $r$  and every  $s$  and  $h$ ,  $\hat{\sigma}^{exp}(a \mid s, h; \zeta_r) > 0$  if and only if  $\kappa(a, h \mid \zeta_r) < \delta_r$  where  $\kappa(a, h \mid \zeta_r)$  is the number of times the player associated with  $\hat{\sigma}$  chose action  $a$  at history  $h$  along  $\zeta_r$ ,

(v) a mapping  $\hat{\alpha}: \bigcup_{r=1}^{\infty} \mathcal{Z}_r \times H \rightarrow [0, 1]$  s.t.  $\hat{\alpha}(\zeta_r, h) < 1$  only if  $\kappa(a, h \mid \zeta_r) < \delta_r$  for some  $a \in A$

and if for each  $r = 1, 2, \dots$ ,  $\zeta_r \in \mathcal{Z}_r$ ,  $s$  and  $h \in H$

$$\hat{\sigma}(s, h \mid \zeta_r) = \hat{\alpha}(\zeta_r, h) * \hat{\sigma}^{opt}(s, h \mid \zeta_r) + [1 - \hat{\alpha}(\zeta_r, h)] * \hat{\sigma}^{exp}(s, h \mid \zeta_r).$$

An *individual learning model* for the social learning stage game is an array of assessment rules and strategic responses, one each for each player. It is *conforming*, if each player's assessment rule is asymptotically empirical and each player's strategic response is asymptotically myopic with calendar-time limitations on experimentation with respect to the assessment rule.

**Definition 7.** A strategy profile  $\sigma^{**}$  is **locally stable**, if there exists some conforming learning model, such that  $\mathbf{P}(\lim_{r \rightarrow \infty} \hat{\sigma}^{opt}(\zeta_r) = \sigma^{**}) > 0$  where  $\hat{\sigma}^{opt}$  denotes the non-experimental part of the array of strategic responses.

**Proposition 9.** In the social learning game independent of feedback and private belief support a strategy profile is locally stable if and only if it is iteratively undominated.

*Proof.* First a similar argumentation as in the proof for the more basic learning process applies, i.e. every player infinitely often observes the complete sequence of choices together with the state of Nature whatever the specification of feedback and private belief support. We thus concentrate on the case with complete feedback on choices, unconditional feedback on the state and homogeneous private belief support.

The proof employs the same inductive argument as for the more basic case: In the first period no inferences about others are necessary and thus assessments are rational by definition. Asymptotic myopia wCTLE then guarantees that players' strategies eventually play arbitrarily like some iteratively undominated strategy  $\sigma^*$  in period 1. We can then employ Lemma D.1 and the SLLNCE to show that the empirical assessment  $\bar{\varphi}_2(\zeta_r)$  for period 2 becomes  $\epsilon$ -close to the rational assessments for this period almost surely eventually. It is then straightforward to show that asymptotic empiricism of general assessment rules and asymptotic myopia with calendar-time lim. on exp. imply that eventually players must play arbitrarily like some iteratively undominated strategy in period 2 as well. This argumentation can obviously be extended inductively to all periods  $i > 2$ .

To see finally that any iteratively undominated strategy can be locally stable (notice that besides in non-generic settings these strategies differ only at histories reached with probability zero) notice that a similar argument can be employed as in Fudenberg and Kreps (1993, Proposition 6.3) which

constructs explicitly an asymptotically empirical assessment rule and an associated myopic strategic response. The interested reader is referred to the paper above or chapter 2 of March (2010). □

As a final remark notice that the Proposition (as its counterpart in the main part of the text) shows that *any* iteratively undominated strategy profile might arise as the outcome of an adaptive process. Indeed going beyond this result towards selecting a unique strategy profile seems to be challenging. In fact even many refined equilibrium concepts (perfect, sequential) do not yield a significantly smaller set of strategy profiles. However under mild additional conditions a unique strategy profile is selected in the limit of a sequence of regular quantal response equilibria (QRE) as payoff disturbances approach zero. We conjecture that this strategy profile may be selected by an adaptive process satisfying some mild additional conditions. Notice that this selection must uniquely define behavior at histories reached with probability zero. While adaptation takes place such histories may occur either because players best respond to mistaken beliefs or because players experiment. However, in either case the behavior is least costly if in line with a player's private belief. Hence, it seems likely that choices while occurring with probability zero in the limit reveal a player's private belief in the medium run – which is the property of limit QRE.



## B The Gaussian-Quadratic Model

In social learning settings more closely related to the classical model of learning in rational expectations the following setting has become very popular (see for instance Vives, 1993). The state of Nature  $\tilde{\theta}$  is normally distributed with mean  $\mu(\theta)$  and variance  $\sigma^2(\theta)$ . Furthermore for realized state  $\theta$  private signals are given by  $\tilde{s}_i = \theta + \tilde{\epsilon}$  where  $\tilde{\epsilon}$  is a noise independent of  $\theta$  and normally distributed around mean 0 with variance  $\sigma_\epsilon^2 > 0$ . This information structure usually is combined with action space  $A = [0, 1]$  and the quadratic payoff (loss) function  $u(a, \theta) = -(a - \theta)^2$ .

The beauty of the Gaussian model stems from the fact that conditional on a signal realization  $s$  by Bayes'rule  $\tilde{\theta}$  is still normally distributed with updated mean and variance given by

$$\begin{aligned}\sigma^2(\theta | s) &= \frac{\sigma^2(\theta) \sigma_\epsilon^2}{\sigma^2(\theta) + \sigma_\epsilon^2}, \\ \mu(\theta | s) &= \frac{\sigma^2(\theta)}{\sigma^2(\theta) + \sigma_\epsilon^2} s + \frac{\sigma_\epsilon^2}{\sigma^2(\theta) + \sigma_\epsilon^2} \mu(\theta).\end{aligned}$$

This follows from the Bayesian formula  $f(\theta | s) = f(\theta) g_\theta(s) / \int_{\theta'} f(\theta') g_{\theta'}(s) d\theta'$  by straightforward calculations. Furthermore expected payoff is maximized by setting  $a(s) = \mu(\theta | s)$ . Hence, if  $\mu(\theta)$ ,  $\sigma^2(\theta)$  and  $\sigma_\epsilon^2$  are known  $a$  perfectly reveals  $s$  and  $f(\theta | a) = f(\theta | s(a))$  where  $s(a)$  is uniquely defined.

Assume as above that players attempt to learn  $f(\theta | a)$  for each  $a \in A$  in a statistical way.<sup>55</sup> That is players assume that conditional on  $a$ ,  $\tilde{\theta}$  is still normally distributed<sup>56</sup> and attempt at determining  $\mu(\theta | a)$  and  $\sigma^2(\theta | a)$  via the sample mean  $\bar{\theta}(a) = \frac{1}{\kappa(a)} \sum_{r: a \in y(r)} \theta(r)$  and the sample variance  $S^2(\theta | a) = \frac{1}{\kappa(a)-1} \sum_{r: a \in y(r)} [\theta(r) - \bar{\theta}(a)]^2$  respectively. Then  $\bar{\theta}(a)$  is normally distributed around mean  $\mu(\theta | a)$  with variance  $\sigma^2(\theta | a)/\kappa(a)$  and  $S^2(\theta | a)$  is asymptotically normally distributed around  $\sigma^2(\theta | a)$  with variance  $2\sigma^4(\theta | a)/\sqrt{n}$ . Furthermore by Cochran's Theorem (Cochran, 1934)  $\bar{\theta}(a)$  and  $S^2(\theta | a)$  are independent. Using these measures for determining mean and variance of the updated normal distribution of  $\tilde{\theta}$  conditional on private signal  $s$  and observation  $a$  players arrive at

$$\begin{aligned}\hat{\mu}(\theta | s, a) &= \frac{S^2(\theta | a)}{S^2(\theta | a) + \sigma_\epsilon^2} s + \left[1 - \frac{S^2(\theta | a)}{S^2(\theta | a) + \sigma_\epsilon^2}\right] \bar{\theta}(a) \\ \text{and } \hat{\sigma}^2(\theta | s, a) &= \frac{\sigma_\epsilon^2 S^2(\theta | a)}{\sigma_\epsilon^2 + S^2(\theta | a)}.\end{aligned}$$

The mapping  $\alpha(S) = \frac{S}{\sigma_\epsilon^2 + S}$  is strictly concave in  $S$ . Therefore by Jensen's inequality  $E[\alpha(S^2(\theta | a))] \leq \alpha(E[S^2(\theta | a)]) = \alpha(\sigma^2(\theta | a))$  provided  $\kappa(a) < \infty$ . Independence of  $\bar{\theta}(a)$  and  $S^2(\theta | a)$  then implies that  $\hat{\mu}(\theta | s, a)$  is adjusted too little towards the signal  $s$ .

We adapt the specific model of heterogeneous belief updating studied in March and Ziegelmeyer (2009) to this setting. That is

$$f_\beta(\theta | s) = \frac{[g_\theta(s)]^\beta f(s)}{\int_{\theta'} [g_{\theta'}(s)]^\beta f(\theta') d\theta'}$$

for some  $\beta > 0$ . Then the posterior distribution of  $\tilde{\theta}$  given signal  $s$  is given by a normal distribution

<sup>55</sup>Clearly, since  $A$  is an interval this requires players to learn an uncountably infinite number of values. However, little is lost by assuming that players partition  $A$  into a finite set of intervals  $A_j$  and learn  $f(\theta | A_j)$  for each  $j$ . If the partition is sufficiently fine since  $\mu(\theta | s(a))$  is continuous in  $a$ ,  $f(\theta | A_j)$  approximates  $f(\theta | a)$  for each  $a \in A_j$ .

<sup>56</sup>Equivalently, they could apply an appropriate statistical test which will be confirmed eventually.

with biased mean and variance

$$\begin{aligned}\hat{\mu}_\beta(\theta | s, a) &= \frac{\beta S^2(\theta | a)}{\sigma_\epsilon^2 + \beta S^2(\theta | a)} s + \left[ 1 - \frac{\beta S^2(\theta | a)}{\sigma_\epsilon^2 + \beta S^2(\theta | a)} \right] \bar{\theta}(a), \\ \sigma_\beta^2(\theta | s, a) &= \frac{\sigma_\epsilon^2 S^2(\theta | a)}{\sigma_\epsilon^2 + \beta S^2(\theta | a)}.\end{aligned}$$

Hence, if  $\kappa(a) < \infty$  there exists  $\beta > 1$  such that  $E[\hat{\mu}_\beta(\theta | s, a)] = \mu(\theta | s, a)$ .

Finally, we show that with the quadratic payoff function this bias in posteriors implies non-optimal choices of Bayesians and therefore fitness benefits of overweighters. Notice first that expected payoff  $U(a | s, x) = E_{\tilde{\theta}}[u(a, \theta) | s, x]$  where  $x$  is the observed action is strictly concave in  $a$ . Thus, it holds  $E_{\tilde{\epsilon}}[U(a | s, x)] < U(E_{\tilde{\epsilon}}[a(s, x)] | s, x)$ . Furthermore  $U(a | s, x)$  is maximized at  $a^* = E[\tilde{\theta} | s, x] = \mu(\theta | s, x)$ . The result above straightforwardly implies that  $E_{\tilde{\epsilon}}[a_1(s, x)] = E_{\tilde{\epsilon}}[\hat{\mu}_1(\theta | s, x)] \neq \mu(\theta | s, x)$  while for some  $\beta > 1$   $E_{\tilde{\epsilon}}[a_\beta(s, x)] = \mu(\theta | s, x)$ . Therefore  $U(E_{\tilde{\epsilon}}[a_1(s, x)] | s, x) < U(E_{\tilde{\epsilon}}[a_\beta(s, x)] | s, x)$ .

## C Limited Learning Opportunities and Non-Bayesian Strategies

In this appendix we investigate in greater detail the relationship between limited learning opportunities and the advantage of non-Bayesian rational behavior. As shown in the main part of the paper (Propositions 5 and 7) generically under limited learning opportunities the adaptive process does not lead Bayesian rational individuals towards adopting the optimal (benchmark) strategy. This implies that Bayesian rationality itself is subject to challenge by alternative ways of responding to information gathered in the adaptive process. In particular experimentation or evolutionary forces may lead individuals to non-Bayesian strategies which ultimately perform better. Our aim in this appendix is to characterize and interpret such alternative strategies.

### C.1 Preliminaries

#### Strategy Space

Apparently, one possible approach would be to study more closely the benchmark strategies derived in Lemmas 2 and 3. Yet, while we can compute and picture these strategies for specific social learning setups the helpfulness of such an illustration is questionable. On the one hand the interpretation of these illustrations is highly ambiguous. More importantly we doubt the potential of the benchmark strategies to capture behavior in real-world environments. Indeed in the main part of the text we argue that non-Bayesian rational behavior which is optimal in field environments with high uncertainty and limited learning opportunities may appear in simpler (laboratory) settings where it leads to systematic errors because individuals do not adapt to each decision situation separately. In other words we rely on the idea that individuals are rule rational and evolve “rules of thumb” (Aumann, 1997) which work well in general. This seems contrary to the assertion that players will follow benchmark strategies which are highly adjusted to specific social learning settings.<sup>57</sup>

The approach we pursue is therefore to investigate the performance of a restricted class of simple strategies. This class derives from an extended model of social learning suggested recently (March and Ziegelmeyer, 2009) to reconcile theoretical predictions with experimental evidence. Indeed a stable finding in the experimental literature on social learning is the tendency of players to rely too strongly upon their private information. The strategies we consider differ in the weight players put on their private information in relation to the information derived from observed actions when forming beliefs. More precisely strategies are parametrized by  $\beta > 0$  and characterized as follows: Given private belief  $s$ , history  $h$ , and assessments  $\varphi$  players form log-likelihood ratios (LLR) according to

$$\ell(s, h \mid \varphi, \beta) = \beta \log \left( \frac{s}{1-s} \right) + \log \left( \frac{\varphi(h \mid 1)}{\varphi(h_i \mid 0)} \right) \quad (7)$$

and invest (reject) if  $\ell(s, h \mid \varphi, \beta) > 0$  ( $\ell(s, h \mid \varphi, \beta) < 0$ ). Accordingly,  $\beta$  is a private information weight. A player overweights private information if  $\beta > 1$ , underweights private information if  $\beta < 1$  and forms beliefs in a Bayesian way if  $\beta = 1$ . The class of strategies hence contains the Bayesian rational strategy as a special case. Our approach therefore permits us to relate our results to both the prevalent Bayesian rational perspective in the theoretical literature as well as the experimental findings.

---

<sup>57</sup>See also Samuelson (2004) who argues that Nature facing prohibitively large costs of enhancing cognitive powers may resort to information processing “short-cuts”.

## Evolutionary Approach

We consider an evolutionary setup where the population is characterized by a distribution  $W$  of simple  $\beta$ -strategies (a distribution on the positive real axis). As shown by March and Ziegelmeyer (2009) the behavior in an extended social learning game with a given distribution of  $\beta$ s is equivalent to behavior in some standard social learning game with a distorted distribution of private information. Therefore our results on the limiting outcomes of the adaptive process straightforwardly extend to this setup. We assume that adaptation has taken place and players' assessments are given by the limiting outcome of the respective adaptive process (the long run outcome in the case of learning across games respectively the outcome after a fixed finite number of repetitions). Under this assumption expected payoffs of "strategy"  $\beta$  given population  $W$  can be straightforwardly derived under the two restrictions of learning opportunities. In particular when learning takes place across games a player of type  $\beta$  invests provided  $[s/(1-s)]^\beta > (\sum_k \pi_k \varphi^k(h|0)) / (\sum_k \pi_k \varphi^k(h|1))$ . Accordingly her expected payoff is given by

$$U_{\Gamma, \pi}(\beta | W) = \sum_{k=1}^K \sum_{i=1}^n \sum_{h_i \in H_i} \frac{\pi_k}{n} \left\{ \varphi^k(h_i | 1, W) [1 - G_1^k(c_\beta(h_i | W))] - \varphi^k(h_i | 0, W) [1 - G_0^k(c_\beta(h_i | W))] \right\} \quad (8)$$

where for each  $k$  and  $\theta$   $\varphi^k(h_1 | \theta, W) = 1$  and

$$\varphi^k(h, a | \theta, W) = \varphi^k(h | \theta, W) * \begin{cases} \int_{\beta \in \text{supp}(W)} [1 - G_\theta^k(c_\beta(h | W))] W(d\beta) & \text{if } a = 1 \\ \int_{\beta \in \text{supp}(W)} G_\theta^k(c_\beta(h | W)) W(d\beta) & \text{if } a = 0 \end{cases}$$

for each  $h$  and  $c_\beta(h | W) = 1 / \left[ 1 + \left( \frac{\sum_k \pi_k \varphi^k(h|1, W)}{\sum_k \pi_k \varphi^k(h|0, W)} \right)^{1/\beta} \right]$  for each  $h$ . On the other hand when the learning horizon is finite given private belief  $s$  and perceived public belief  $b'(\emptyset, h) = b(\emptyset, h) + \tilde{\epsilon}_h$  a player of type  $\beta$  invests if  $[s/(1-s)]^\beta > (1 - b(\emptyset, h) - \tilde{\epsilon}_h) / (b(\emptyset, h) + \tilde{\epsilon}_h)$  or equivalently if  $\tilde{\epsilon}_h > \frac{(1-s)^\beta}{s^\beta + (1-s)^\beta} - b(\emptyset, h)$ . Accordingly, her expected payoff is given by

$$U_\eta(\beta | W) = \frac{1}{n} \sum_{i=1}^n \int_{\underline{b}}^{\bar{b}} \sum_{h_i \in H_i} [1 - \Phi_h(\chi_\beta(s, h | W))] * [\varphi_\eta(h_i | 1, W) dG_1(s) - \varphi_\eta(h_i | 0, W) dG_0(s)] \quad (9)$$

where for each  $\theta$   $\varphi_\eta(h_1 | \theta, W) = 1$  and

$$\varphi_\eta(h, a | \theta, W) = \varphi_\eta(h | \theta, W) * \begin{cases} \int_{\beta \in \text{supp}(W)} \int_{\underline{b}}^{\bar{b}} [1 - \Phi_h(\chi_\beta(s, h | W))] dG_\theta(s) W(d\beta) & \text{if } a = 1 \\ \int_{\beta \in \text{supp}(W)} \int_{\underline{b}}^{\bar{b}} \Phi_h(\chi_\beta(s, h | W)) dG_\theta(s) W(d\beta) & \text{if } a = 0 \end{cases}$$

for each  $h$ ,  $\chi_\beta(s, h | W) = \frac{(1-s)^\beta}{s^\beta + (1-s)^\beta} - \frac{\varphi_\eta(h|1, W)}{\varphi_\eta(h|1, W) + \varphi_\eta(h|0, W)}$  for each  $s$  and  $h$ , and  $\Phi_h$  denotes the cdf of the normal distribution with mean zero and standard deviation  $\eta_h$ .

We have in mind an evolutionary process which selects strategies according to their relative fitness where fitness of strategies is given by the above (ex-ante) expected payoffs. However, given the complexity of the fitness expressions an exact theoretical analysis of the limiting outcome of such a process is beyond the scope of this addendum. In contrast the simple recursive structure allows an easy numerical computation. We therefore rely on numerical methods to illustrate the limiting outcomes. More precisely we examine the relationship between the social learning environment and the surviving strategies with the help of an implementation of the discrete-time replicator dynamics. In order to do

so we first discretize the strategy space (the positive real axis). The reduced strategy set is given by

$$\mathcal{B} = \{0.14, 0.17, 0.20, 0.25, 0.30, 0.37, 0.45, 0.55, 0.67, 0.82, 1.00, 1.22, 1.49, 1.82, 2.23, 2.72, 3.32, 4.06, 4.95, 6.05, 7.39\}.^{58}$$
(10)

Let  $W_t(\beta)$  denote the relative frequency of type  $\beta$  after  $t$  steps. We assume that type frequencies evolve according to the replicator equation

$$W_{t+1}(\beta) = \frac{1 + \Delta W_t(\beta) U(\beta | W_t)}{1 + \Delta \sum_{\beta \in B} [W_t(\beta) * U(\beta | W_t)]}.^{59}$$
(11)

Here  $\Delta$  denotes the step size, i.e. length of the time interval. Given the size of the payoffs we will fix it at  $\Delta = 100$  in order to obtain convergence of the system within a reasonable number of steps. Finally, we consider as starting distribution  $W_0$  the uniform distribution on  $B$  which puts all strategies on an equal footing.

### Social Learning Setting

The space of possible distributions of private information is huge and many stories can be told to justify either. We stick with two extreme variants which have been frequently assumed in the literature: The *symmetric binary signal (SBS) distribution* and the *Gaussian signal distribution*.<sup>60</sup> The binary signal distribution is characterized by a binary signal space  $S = \{0, 1\}$  and conditional probabilities  $Pr(\tilde{s} = 1 | \tilde{\theta} = 1) = Pr(\tilde{s} = 0 | \tilde{\theta} = 0) = q$  where  $1/2 < q < 1$  denotes the *signal precision*. On the other hand the Gaussian signal distribution is given by signal space  $S = \mathbb{R}$  such that conditional on  $\theta$   $\tilde{s}$  is normally distributed around mean  $\theta$  with precision  $\rho > 0$  (respectively standard deviation  $1/\sqrt{\rho}$ ). The main advantage of these distributions is their dependence upon a single easily interpretable parameter. In addition the associated private log-likelihood ratio  $\ell(\tilde{s}, \emptyset) = \log\left(\frac{\tilde{s}}{1-\tilde{s}}\right)$  follows a shifted Bernoulli (conditional on one of the states) respectively Gaussian distribution, something that will become important in the context of learning across games. Finally, with these two distributions we consider both a setting with bounded and one with unbounded private beliefs which enables us to investigate the dependence of the results upon this crucial property.

### Presentation of Results

Results will be presented in the form of arrays of bar charts. Each array will be identified with a specific underlying distribution of private signals and a specific sequence length  $n$  (and in addition in the case of adaptation across games the number of signals in one of the two basic social learning games). Furthermore in each array the parameter of the underlying private signal distribution ( $q$  respectively  $\rho$ ) varies with the columns of the array and a parameter specific to the limit on learning opportunities (the number of private signals  $L_2$  respectively the learning horizon  $R$ ) varies with the rows of the array. Finally, each bar chart depicts the distribution  $W_T$  after a fixed number of steps  $T$ . Figure 1 shows an example of such a chart.

Several properties are noteworthy. First, bars are sorted from left to right by increasing  $\beta$  (on the set  $\mathcal{B}$ ). Second, charts are scaled such that the range of the  $y$ -axis is always the complete interval

<sup>59</sup>See for instance Fudenberg and Levine (1998, Section 3.11).

<sup>60</sup>See e.g. Chamley (2004b, Chapter 2).

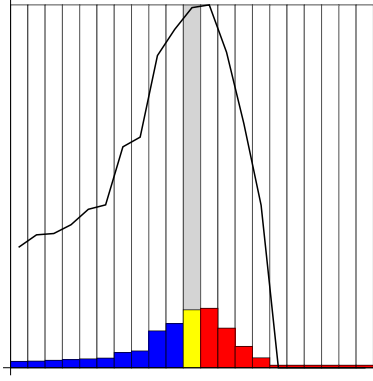


Figure 1: Example chart for the presentation of the numerical results.

$[0, 1]$ . Third, in order to facilitate distinction the range of possible values for each bar is framed. In addition the area for  $\beta = 1$  has a gray background. Fourth, the bars itself are colored differently depending on whether they refer to underweighting of private information ( $\beta < 1$ , BLUE), Bayesian updating ( $\beta = 1$ , YELLOW), or overweighting of private information ( $\beta > 1$ , RED). Finally, each chart contains a graph depicting fitness of each type for the depicted population  $W_T$ . This enables to get an impression about the future evolution of the population in cases where convergence has not yet taken place. The graphs are re-scaled such that lowest fitness is at  $y = 0$  and highest fitness at  $y = 1$ . In cases where all types have exactly the same fitness a straight line is given at  $y = 0.5$ .

We now discuss in turn the two approaches towards limited learning opportunities and present the numerical results.

## C.2 Adaptation across Games

We consider a class of global social learning games given by  $\langle n, A, u, \Theta, (G_0^k, G_1^k)_{k=1}^2, \pi \rangle$  such that  $A = \Theta = \{0, 1\}$ ,  $u(a, \theta) = a * (2\theta - 1)$ , and  $\pi = (1/2, 1/2)$ . Hence, each of these global games is comprised of two basic social learning games, equally likely, which differ only in the distribution of private information. We restrict ourselves further in assuming that the underlying distribution of private signals in both these basic games is the same. The games differ in the number of independent private signals each player receives. Denote the number of signals received in (basic) game  $k \in \{1, 2\}$  by  $L_k$  where w.l.o.g.  $L_1 < L_2$ .<sup>61</sup>

### Symmetric Binary Signals

Consider first an underlying binary signal distribution. Given  $L$  independent draws private beliefs are distributed on the set

$$B_L = \left\{ b_j = \frac{q^j (1-q)^{L-j}}{q^j (1-q)^{L-j} + (1-q)^j q^{L-j}} : j = 0, 1, \dots, L \right\}$$

<sup>61</sup>Again there is a myriad of modeling possibilities even when restricting to a narrow class of private belief distributions and a two-game setting. Ideally one strives to capture properties of real-world environments. Our choice contrasts distributions of private information which arise endogeneously determined only by the number of signals obtained by each player. In a world with costly private signals failures of individuals to acknowledge correlation of (own and others') private information qualities may then be related to imperfect information about budget constraints and search costs.

according to probabilities  $Pr(b_j | \tilde{\theta} = 1) = \binom{L}{j} q^j (1-q)^{K-j}$  and  $Pr(b_j | 0) = \binom{L}{j} (1-q)^j q^{K-j}$ .

Figures 2 and 3 depict the population after  $T = 5000$  steps for different values of  $n$  when the number of signals in game 1 is fixed at  $L_1 = 3$ .<sup>62</sup> In these figures columns capture different precisions  $q$  of private signals while rows capture different numbers of signals  $L_2$  in game 2.

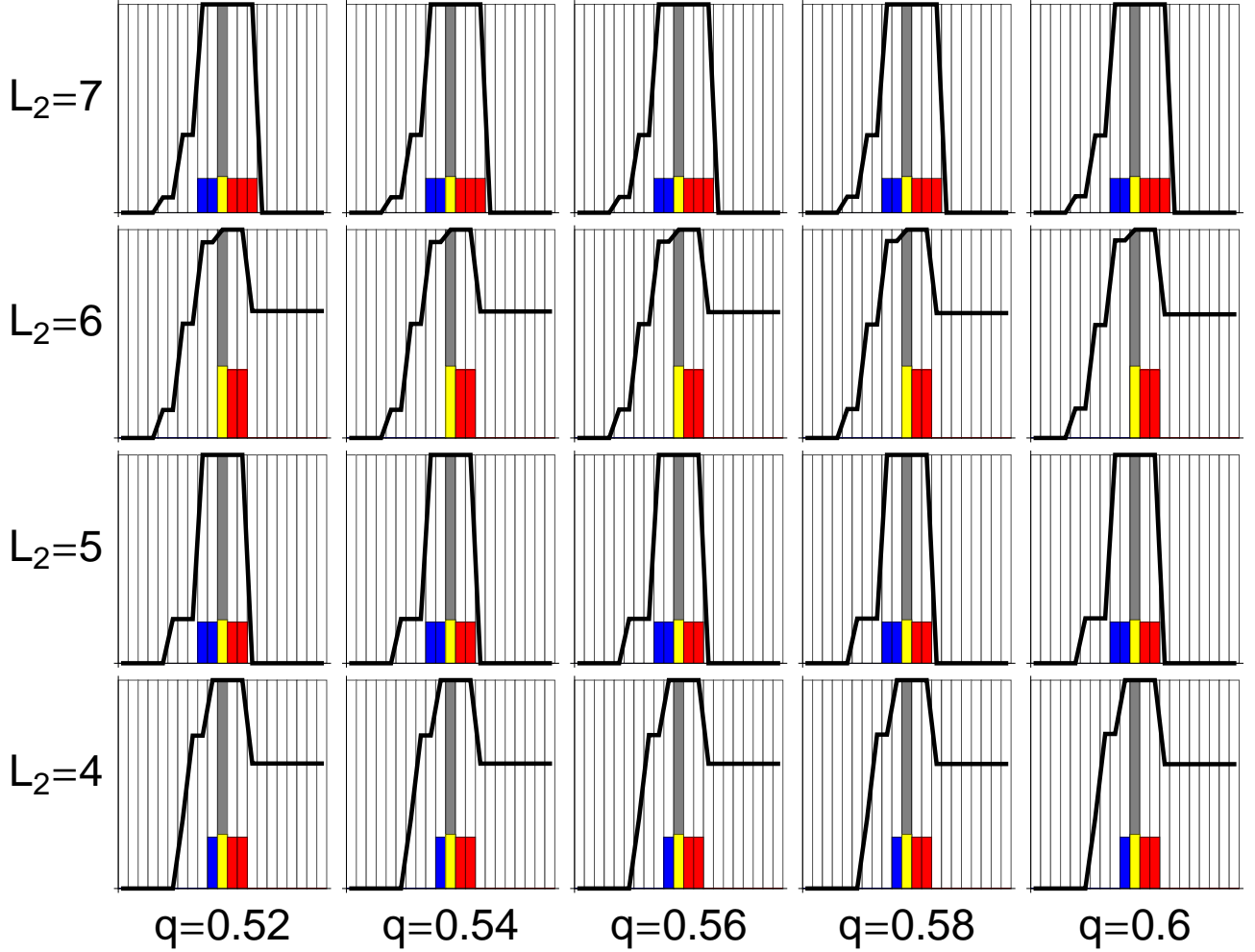


Figure 2: Population after 5000 steps when learning takes place across games, players receive binary signals,  $L_1 = 3$ , and  $n = 2$ .

<sup>62</sup>Similar results are obtained for  $L_1 \in \{2, 4, 5\}$ .

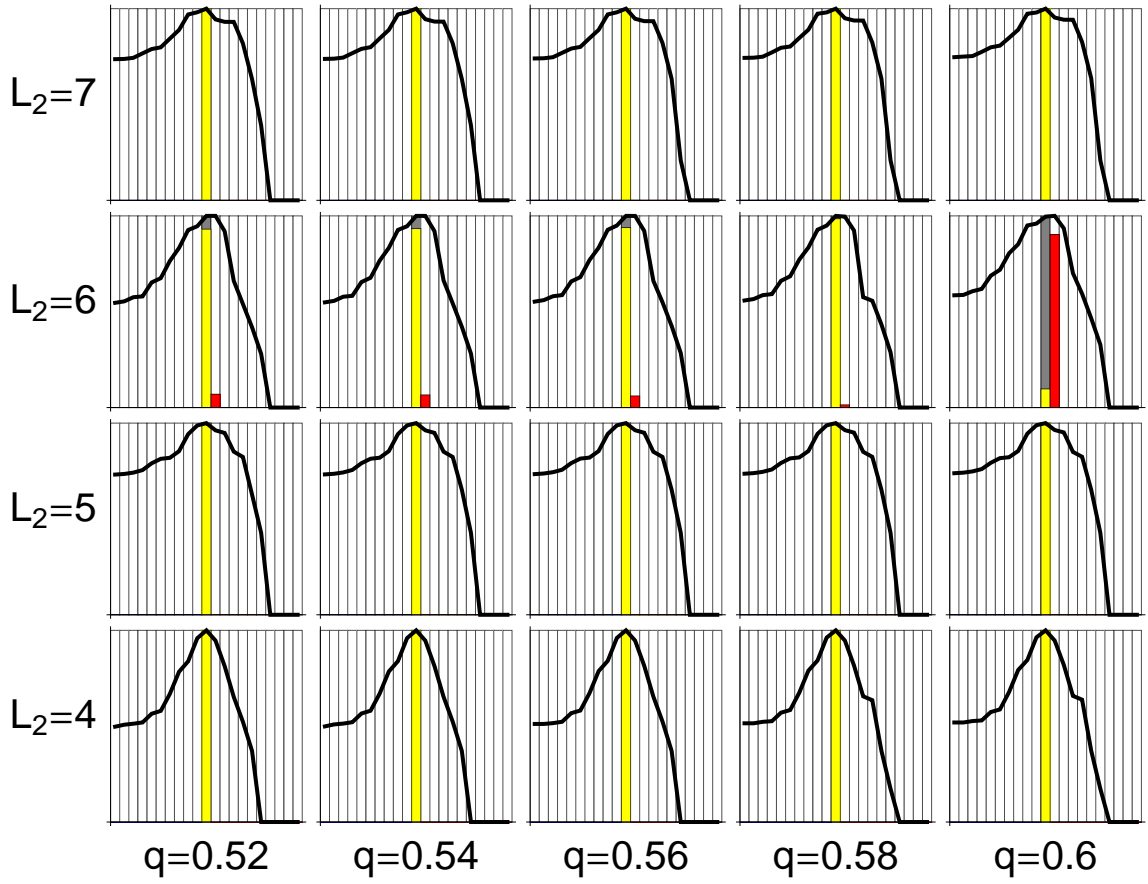
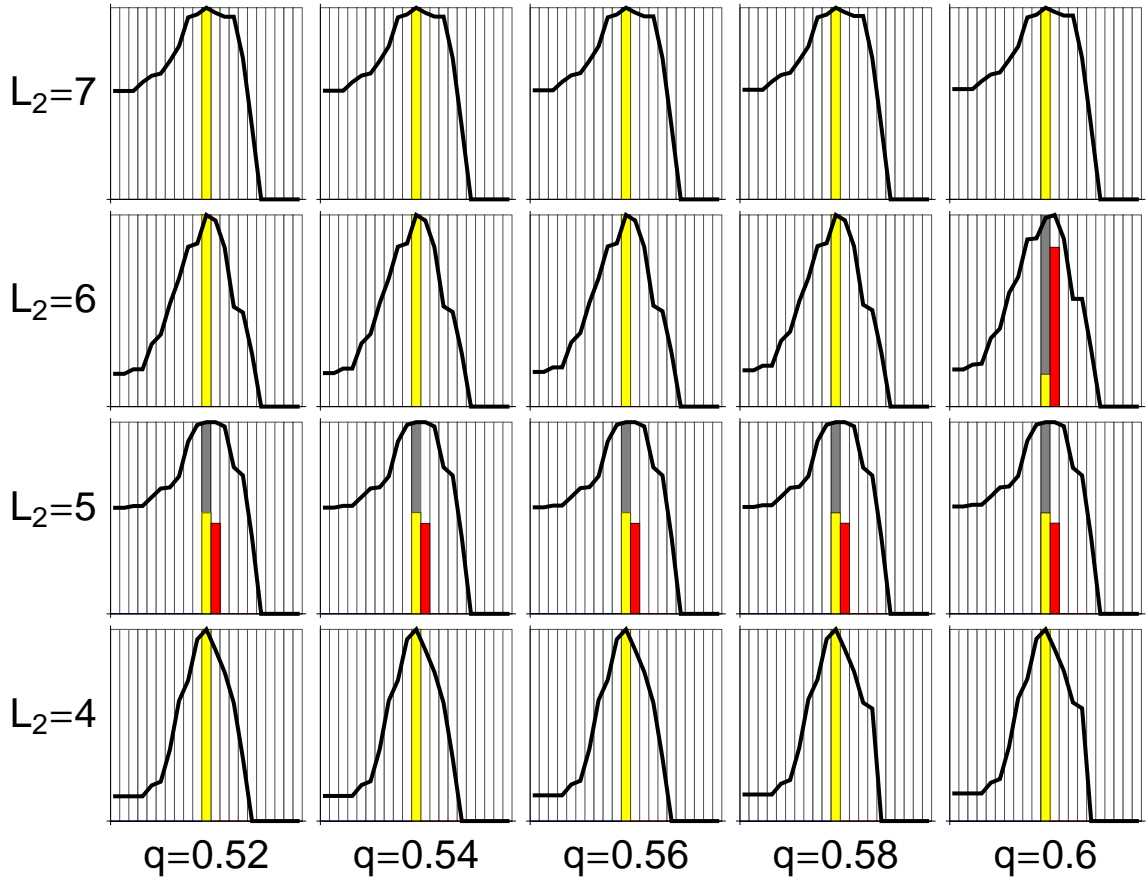


Figure 3: Population after 5000 steps when learning takes place across games, players receive binary signals,  $L_1 = 3$ , and  $n = 4$  (TOP) respectively  $n = 6$  (BOTTOM).



As one can see while for  $n = 2$  both (moderate) over- and underweighting may survive in the evolutionary process, Bayesian updating (at least approximately given the discretization of the strategy space) emerges as the limiting outcome of the replicator dynamics for larger  $n$  in many environments. We now check the robustness of these results in a setting with normally distributed signals.

### Gaussian Signals

As mentioned previously Gaussian private signals imply that the log-likelihood ratio  $\ell(\tilde{s}, \emptyset)$  follows a Gaussian distribution as well. In addition for a sequence  $(s_1, \dots, s_L)$  of  $L$  independent private signals the LLR satisfies  $\ell((s_1, \dots, s_L), \emptyset) = \sum_{j=1}^L \ell(s_j, \emptyset)$ . Since a sum of  $L$  independent normally distributed random variables follows a normal distribution as well, the distribution of the LLR with  $L$  independent private signals is easily calculated.

The population after  $T = 3000$  steps is presented in figures 4 and 5 for various values of  $n$  and  $L_1 = 1$  signals in game 1. As before different columns (rows) of the array capture different signal precisions  $q$  (number of signals  $L_2$  in game 2).

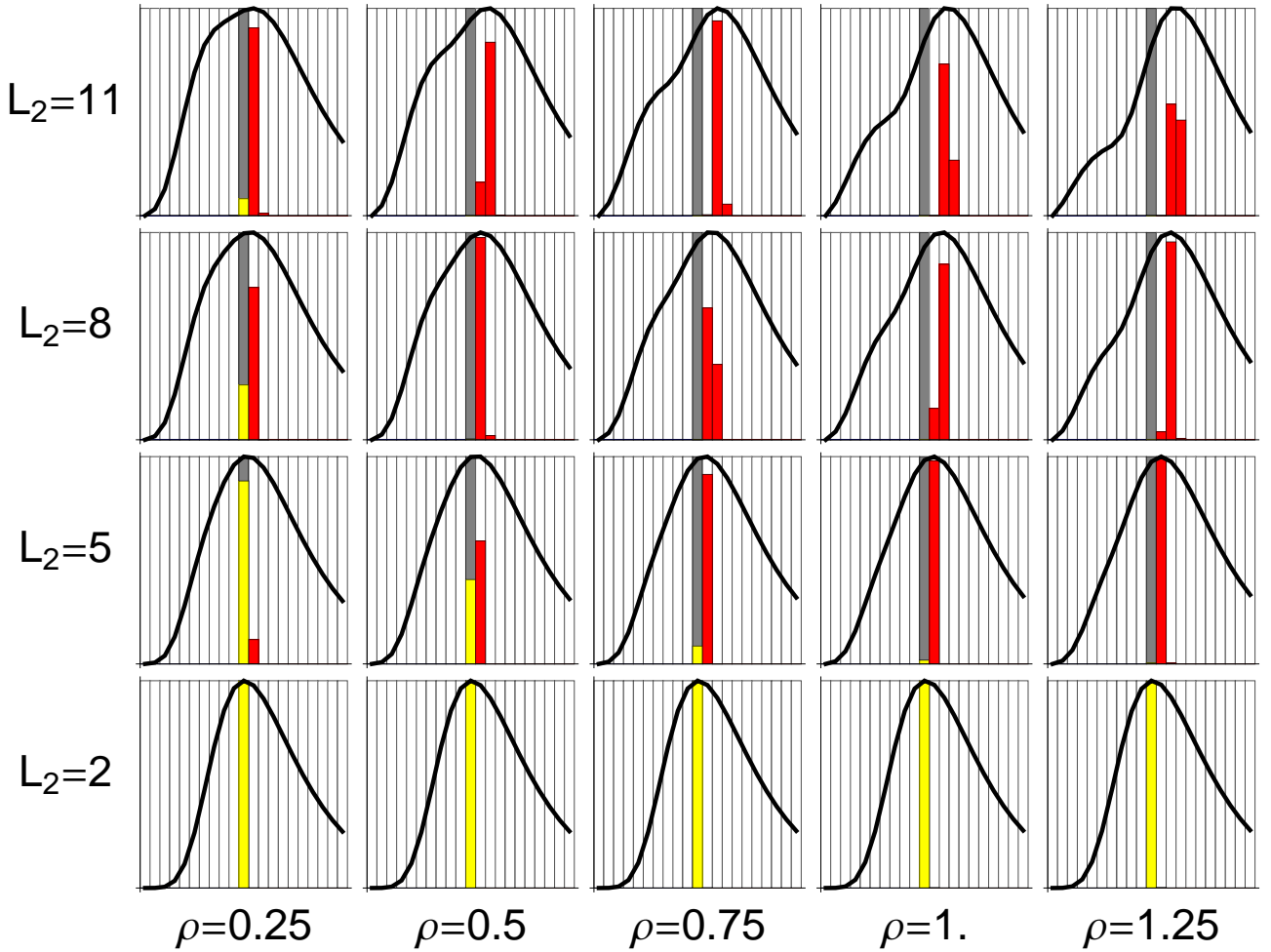


Figure 4: Population after 3000 steps when learning takes place across games, players receive Gaussian signals,  $L_1 = 1$ , and  $n = 2$ .

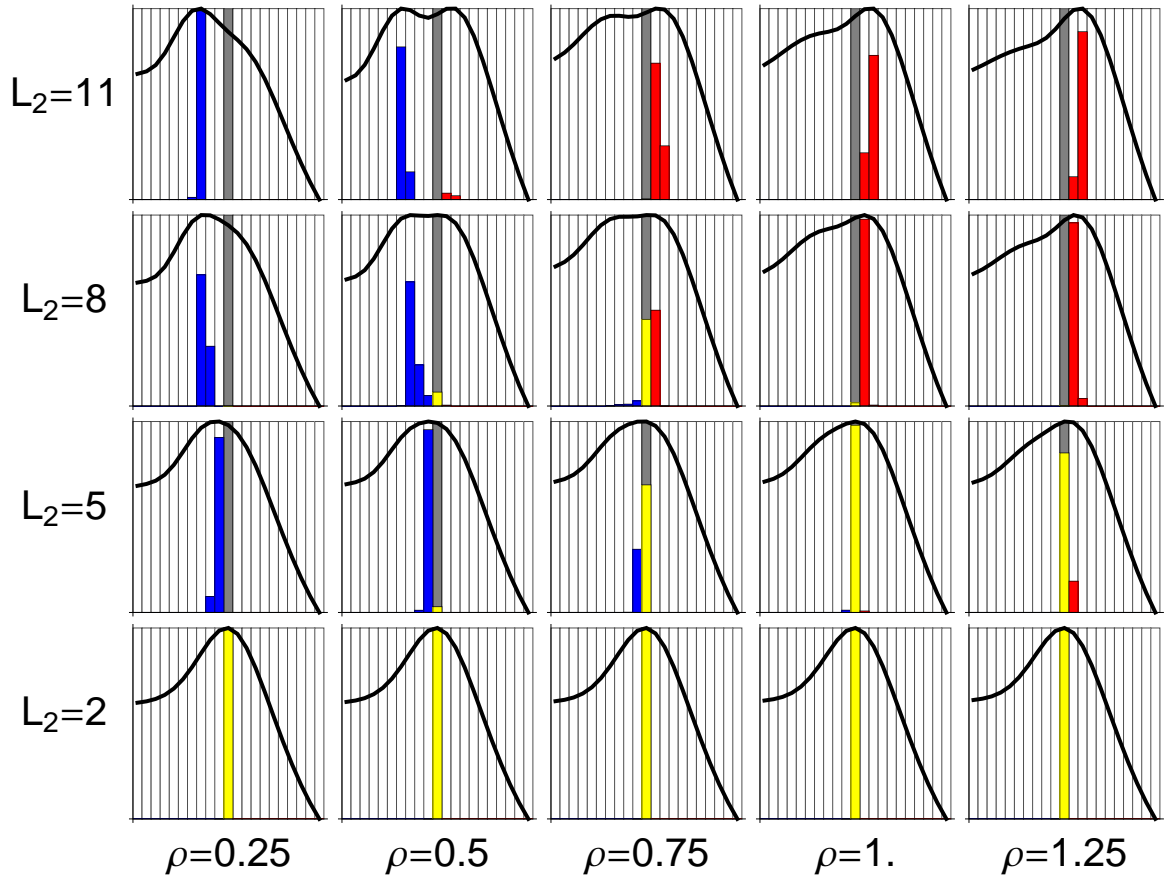
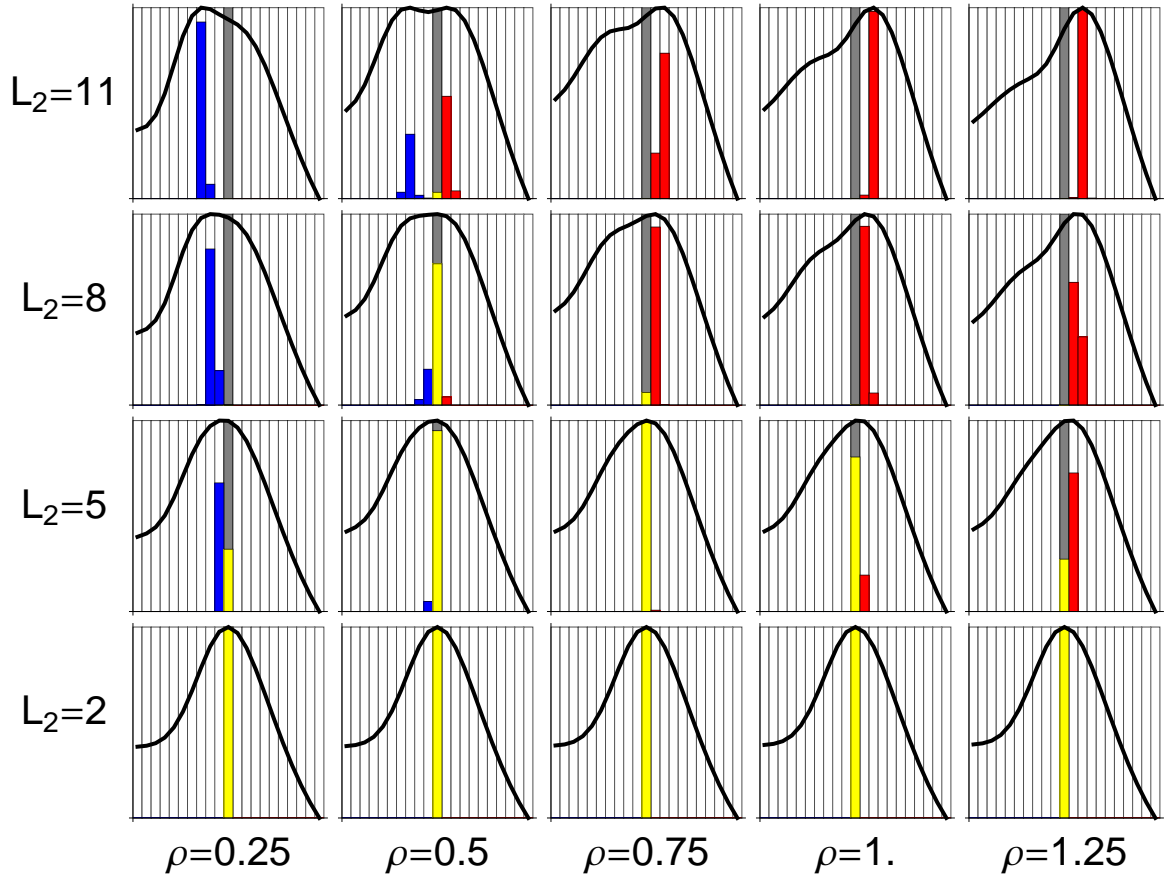


Figure 5: Population after 3000 steps when learning takes place across games, players receive Gaussian signals,  $L_1 = 1$ , and  $n = 4$  (TOP) respectively  $n = 6$  (BOTTOM).

Given the results for both distributions of private signals we may now state the following result.

**Result 3.** *When adaptation takes place across games both moderate over- and underweighting robustly survive in the evolutionary process. In particular moderate overweighers of private information are more likely to be present in a population (i) the shorter the social learning sequence, (ii) the larger the precision of private signals, and (iii) the larger the difference of the basic social learning settings. On the other hand moderate underweighters are more likely to be present, the longer the sequence, the lower the signal precision, and the larger the difference between settings.*

### C.3 Finite Learning Horizon

We now consider the survival of strategies when learning takes place within a finite number of rounds. As argued in the main part of the paper under this condition players' assessments remain noisy and players ultimately play a noisy social learning game  $\langle n, \Theta, A, u, (G_0, G_1), \boldsymbol{\eta} \rangle$  where  $\boldsymbol{\eta} = (\eta_h)_{h \in H}$  with  $\eta_{h_1} = 0$  is the vector of standard deviations of the unbiased, normally distributed noise terms affecting public beliefs. Clearly,  $\boldsymbol{\eta}$  is uniquely determined by the learning horizon  $R$ , the structure of the social learning game, and the specific learning path  $\zeta_R$ . In line with our previous argumentation we make the following assumption

$$\eta_h = \sqrt{\frac{1}{R} \frac{2 \varphi_{\boldsymbol{\eta}}(h \mid 1, W) \varphi_{\boldsymbol{\eta}}(h \mid 0, W)}{[\varphi_{\boldsymbol{\eta}}(h \mid 1, W) + \varphi_{\boldsymbol{\eta}}(h \mid 0, W)]^3}} \quad (12)$$

for each  $h$  where  $\varphi_{\boldsymbol{\eta}}(h \mid \theta, W)$  are the conditional probabilities of histories derived previously. This assumption reflects the idea that players use sample means to estimate public beliefs. These are asymptotically normally distributed around the correct public belief  $b(\emptyset, h)$  with variance  $b(\emptyset, h) [1 - b(\emptyset, h)] / \kappa(h \mid \zeta_R)$  and we approximate the number of occurrences of history  $h$  via the expected number of occurrences, i.e.  $\kappa(h \mid \zeta_R) \approx R * \frac{\varphi_{\boldsymbol{\eta}}(h \mid 1, W) + \varphi_{\boldsymbol{\eta}}(h \mid 0, W)}{2}$ .

Notice that with this assumption we will in general strongly *overestimate* the precision of players' assessments: While players in the first period behave according to the Bayesian rational strategy in each round, players later in the sequence will in general behave suboptimally until their assessments are sufficiently close to their true counterparts. This in turn leads assessments about these players' strategies to be even further off target. In contrast our assumption requires that (noisy) assessments result from  $R$  observations of game play were players are bound to their strategies given these (noisy) assessments. As will be shown assuming less precise assessments would only strengthen our results.

Results are presented in figures 6 to 8. We depict populations after 3000 (resp. 1000) steps for the binary (resp. Gaussian) signal distribution for different sequence lengths  $n$  ( $n = 2, 4, 6$  with binary signals and  $n = 2, 4$  with Gaussian signals), signal precisions  $q$  resp.  $\rho$  (columns of the arrays), and learning horizon  $R$  (rows of the arrays).

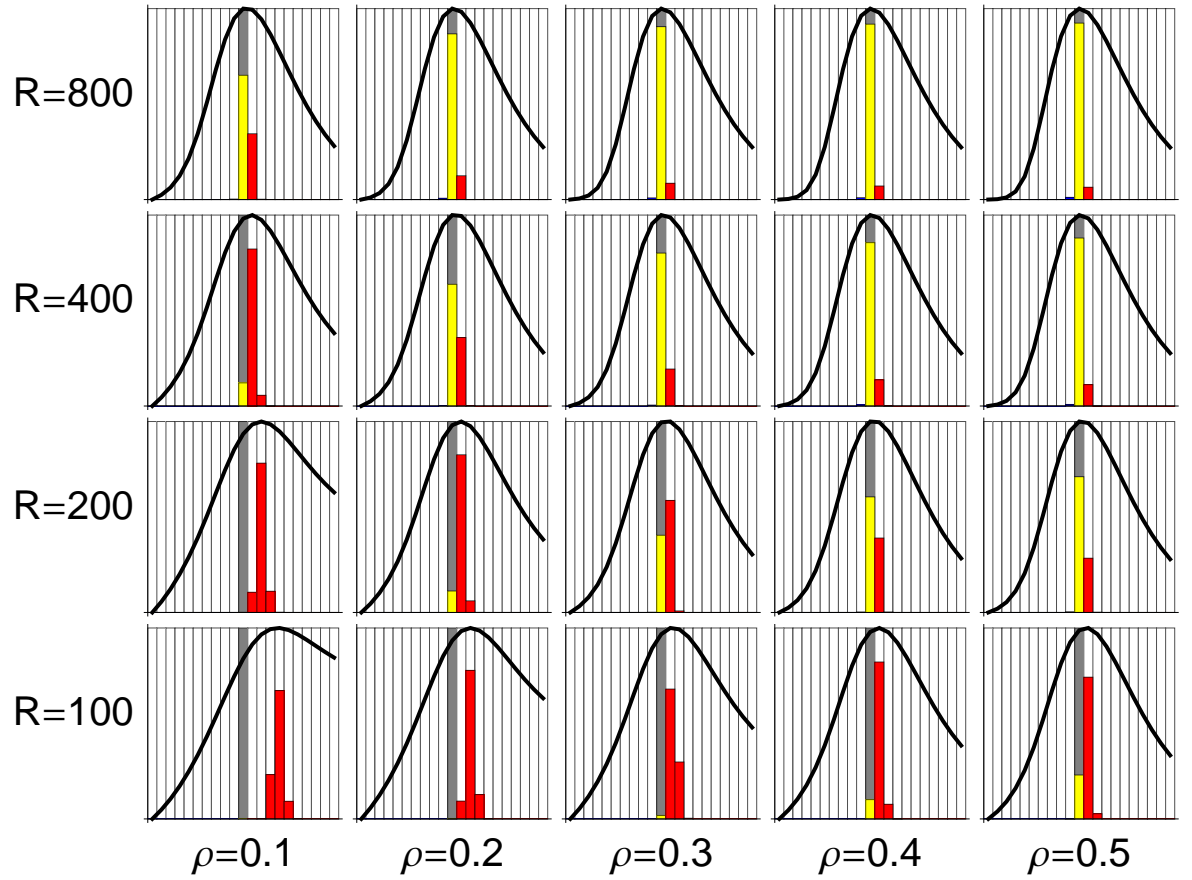
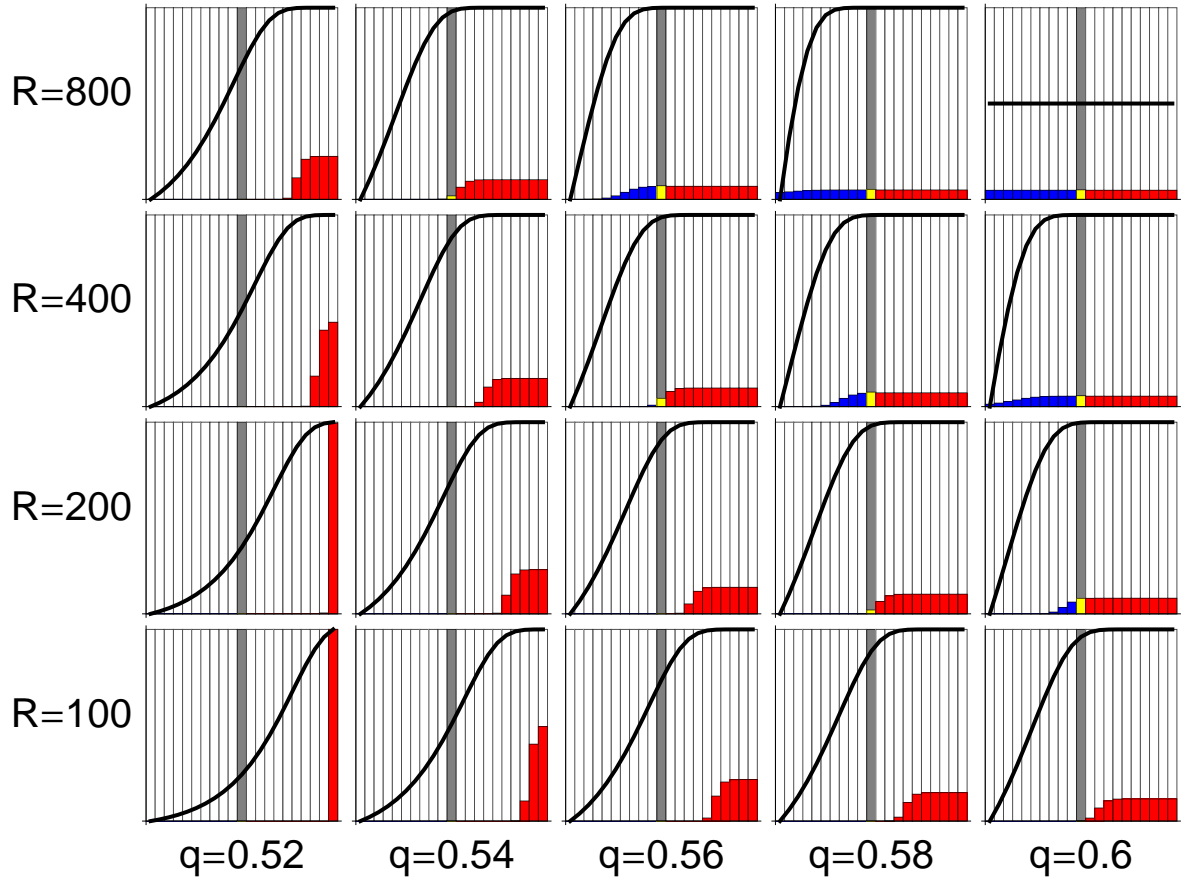


Figure 6: Population after 3000 (TOP) respectively 1000 (BOTTOM) steps when the learning horizon is finite,  $n = 2$ , and players receive binary (TOP) respectively Gaussian (BOTTOM) signals.

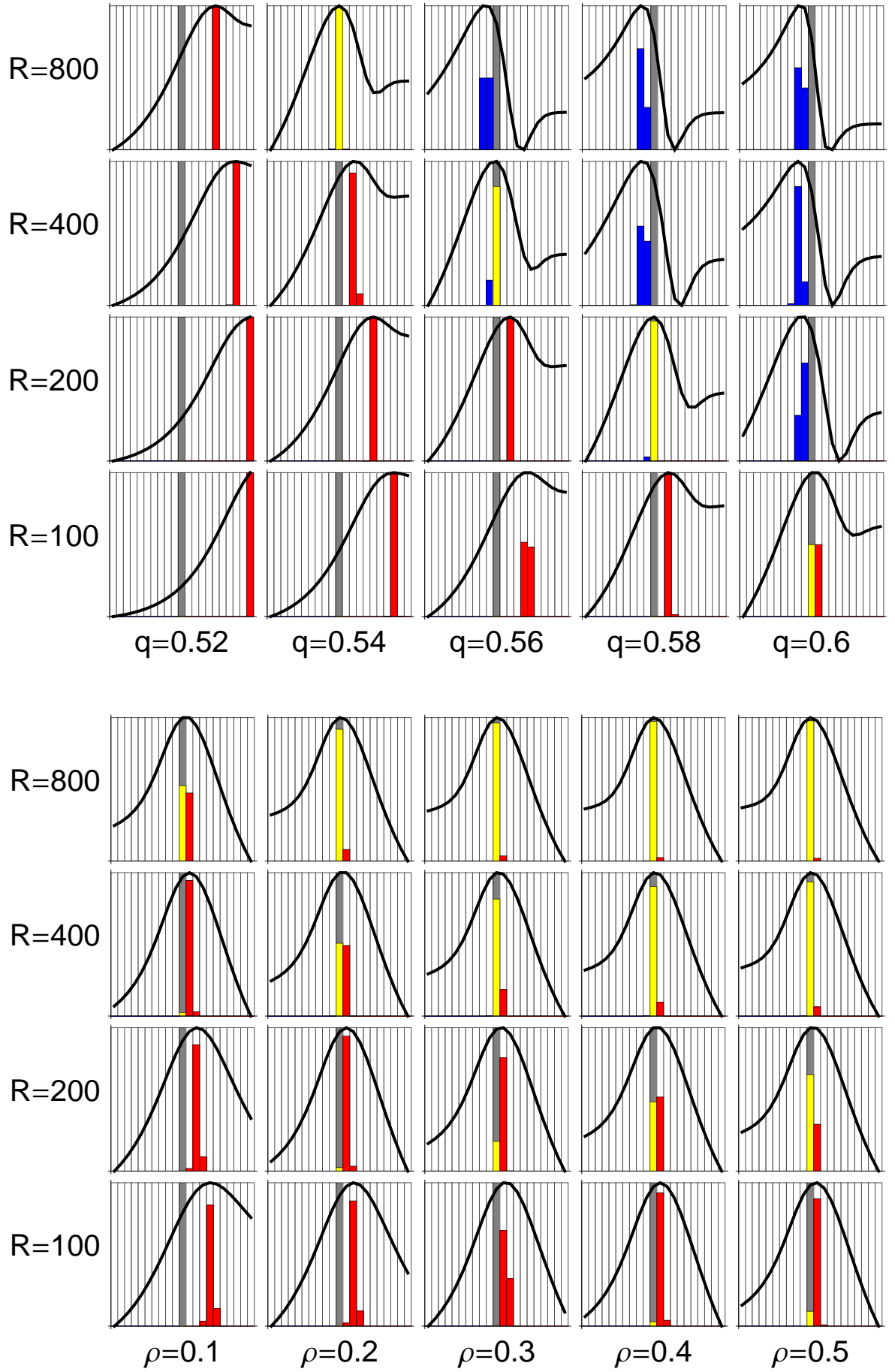


Figure 7: Population after 3000 (TOP) respectively 1000 (BOTTOM) steps when the learning horizon is finite,  $n = 4$ , and players receive binary (TOP) respectively Gaussian (BOTTOM) signals.

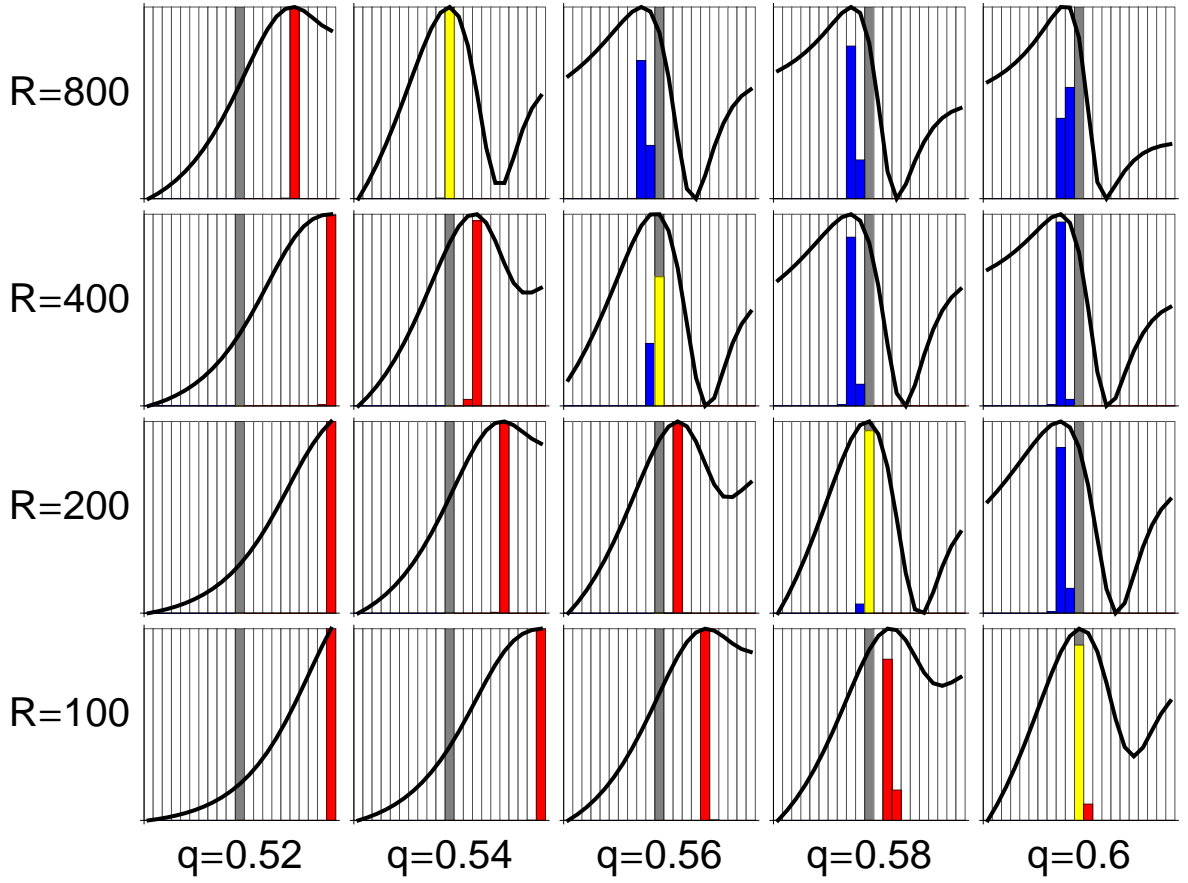


Figure 8: Population after 3000 steps when the learning horizon is finite,  $n = 6$ , and players receive binary signals.

Again we find that both over- and underweighting of private information robustly survives. Remarkably though unlike in the case of adaptation across games overweighting seems to robustly outperform other strategies in settings with low signal precision and short learning horizon even in longer social learning games. Moreover more extreme forms of overweighting survive in such cases. Result 4 below summarizes our findings.

**Result 4.** *If the learning horizon is finite such that players' assessments remain noisy, both over- and underweighting of private information robustly survive in an evolutionary process. In particular settings which involve a shorter learning horizon and a lower precision of private information are more favorable to overweighting while underweighting is favored by a longer learning horizon, higher precision of private information, and longer social learning sequences. In addition a finite learning horizon may favor extreme forms of overweighting.*

## D Omitted Proofs

**Proof of Proposition 1.** We show that all iteratively undominated strategy profiles satisfy conditions (i) and (ii) and barring non-genericities generate the same unique outcome of the social learning game. More precisely we show that all strategy profiles satisfying (i) and (ii) coincide except at histories which occur with probability 0 or at private beliefs which either occur with probability 0 as well or whose occurrence with strictly positive probability is a non-generic property of the social learning context. This in turn implies that all histories occur with the same conditional probability under any iteratively undominated strategy profile proving the claim.

We proceed by induction on the set of players. Notice that due to the sequential structure of the game and the absence of payoff externalities the set of player  $i$ 's rationalizable strategies is determined solely by her beliefs about her predecessors. In particular a strategy of player 1 is undominated if and only if it maximizes

$$U_1(\sigma_1, \sigma_{-1}) = U_1(\sigma_1) = \frac{1}{4} \int_{\underline{b}}^{\bar{b}} \sigma_1(s_1, \emptyset) \frac{2s_1 - 1}{s_1} dG_1(s_1)$$

where we have used that  $dG_0(s)/dG_1(s) = (1-s)/s$ . Hence, any strategy of player 1 which is not dominated must satisfy  $\sigma_1(s_1, \emptyset) = 1$  if  $s_1 > 1/2$  and  $\sigma_1(s_1, \emptyset) = 0$  if  $s_1 < 1/2$ . On the other hand dominance does not restrict player 1's strategy at  $s_1 = 1/2$ . However this case arises with strictly positive probability only if the private belief distributions have an atom at  $1/2$  which is clearly a non-generic property of the social learning setting. Therefore henceforth we assume that  $dG_\theta(1/2) = 0$  for each  $\theta \in \Theta$ . In this case player 1 has a unique strictly dominant strategy.

Let  $i \geq 2$  and assume that for each  $j < i$  each iteratively undominated  $\sigma_j$  satisfies properties (i) and (ii) of the proposition and uniquely determines behavior except for a well-defined finite number of private beliefs. Clearly that the common support of the private belief distributions contains atoms at exactly these private beliefs is a non-generic property of the social learning setting. In any social learning setting which does not have this property any profile  $\sigma_{<i}^*$  of iteratively undominated strategies for players  $j < i$  generates the same distribution of histories  $h_i \in H_i$ . This follows directly from the expression

$$Pr(h_i | \tilde{\theta} = \theta, \sigma_{<i}^*) = \prod_{j < i} \int_{\underline{b}}^{\bar{b}} \sigma_j^*(a_j | s_j, h_j) dG_\theta(s_j)$$

for each  $\theta \in \Theta$  where  $a_j = h_i(j)$  and  $h_j \subset h_i$ . Player  $i$ 's ex-ante expected payoff against an iteratively undominated strategy profile  $\sigma_{<i}^*$  may then be written as

$$U_i(\sigma_i, \sigma_{<i}^*) = \frac{1}{4} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \sigma_i(s_i, h_i) [s_i Pr(h_i | \tilde{\theta} = 1, \sigma_{<i}^*) - (1-s_i) Pr(h_i | \tilde{\theta} = 0, \sigma_{<i}^*)] \frac{1}{s_i} dG_1(s_i)$$

where we have used once more the proportional property of private beliefs. For histories  $h_i$  such that  $Pr(h_i | \tilde{\theta} = \theta, \sigma_{<i}^*) > 0$  for some  $\theta \in \Theta$  (and thus for each  $\theta \in \Theta$  since strategies depend on private beliefs and private belief distributions have common support) it is easy to see that any iteratively undominated strategy  $\sigma_i^*$  of player  $i$  must satisfy  $\sigma_i^*(s_i, h_i) = 1$  if  $s_i > 1 - b^*(\emptyset, h_i)$  and  $\sigma_i^*(s_i, h_i) = 0$

if  $s_i < 1 - b^*(\emptyset, h_i)$  where

$$b^*(\emptyset, h_i) = \frac{Pr(h_i | \tilde{\theta} = 1, \sigma_{< i}^*)}{Pr(h_i | \tilde{\theta} = 1, \sigma_{< i}^*) + Pr(h_i | \tilde{\theta} = 0, \sigma_{< i}^*)}.$$

Furthermore at such histories the player is tied only if the common support of private beliefs has an atom at  $1 - b^*(\emptyset, h_i)$ . Since  $H_i$  is finite this concerns only a finite number of points for player  $i$  and is thus a non-generic property. On the other hand for histories such that  $Pr(h_i | \tilde{\theta} = \theta, \sigma_{< i}^*) = 0$  for each  $\theta \in \Theta$  any assignment  $\sigma^*(s_i, h_i)$  can be part of an iteratively undominated strategy. Notice that properties (i) and (ii) of the proposition do not restrict  $\sigma_i$  in such cases either and are thus trivially satisfied. However, since

$$Pr(h_{i+1} | \tilde{\theta} = \theta, \sigma_{< i}, \sigma_i) = Pr(h_i | \tilde{\theta} = \theta, \sigma_{< i}) * \int_{\underline{b}}^{\bar{b}} \sigma_i(a_i | s_i, h_i) dG_\theta(s_i)$$

$Pr(h_i | \tilde{\theta} = \theta, \sigma_{< i}) = 0$  implies  $Pr(h_{j'} | \tilde{\theta} = \theta, \sigma_{< i}, \sigma_i, \dots, \sigma_{j'-1}) = 0$  for  $h_{j'} \supset h_i$  independent of the behavior following such histories. Thus behavior of player  $i$  at such histories influences neither the outcome of the game nor the iteratively undominated strategies of later players.

In conclusion absent ties iterated strict dominance uniquely characterizes the behavior of players at all histories occurring with strictly positive probability and thus implies a unique outcome of the social learning game. Moreover ties are a non-generic property of a social learning game since they can be ruled out by removing a finite number of points from the support of private beliefs.  $\square$

**Proof of Lemma 1.** We need to show that along almost any (infinite) learning path  $\zeta_\infty$  for each  $i = 1, \dots, n$ , each  $\epsilon > 0$  and each  $R$  player  $i$  eventually  $\epsilon$ -best responds only to strategies played in repetitions  $r > R$ . Here player  $i$   $\epsilon$ -best responds to a set of behavioral strategies  $T \subseteq \Sigma$  if and only if  $\sigma_i \in \hat{\Sigma}_i^\epsilon(T)$  where

$$\hat{\Sigma}_i^\epsilon(T) = \left\{ \sigma_i \in \Sigma_i : \text{for each } \sigma'_i \in \Sigma_i \exists \sigma_{-i} \in T_{-i} \text{ s.t. } U_i(\sigma_i, \sigma_{-i}) + \epsilon > U_i(\sigma'_i, \sigma_{-i}) \right\}.$$

Recall that

$$U_i(\sigma_i, \sigma_{-i}) = \frac{1}{4} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \sigma_i(s_i, h_i) \left[ s_i Pr(h_i | \tilde{\theta} = 1, \sigma_{-i}) - (1 - s_i) Pr(h_i | \tilde{\theta} = 0, \sigma_{-i}) \right] \frac{1}{s_i} dG_1(s_i)$$

where the probabilities  $Pr(h_i | \tilde{\theta} = \theta, \sigma_{-i})$  only depend upon strategies  $\sigma_j$  for  $j < i$ . Fix learning path  $\zeta_\infty$ ,  $\epsilon > 0$  and  $0 < R < \infty$  and let

$$T_{R,r} = \{ \sigma_{-i} : \sigma_{-i} = \hat{\sigma}_{-i}(\zeta_\rho) \text{ for some } R < \rho < r \text{ where } \zeta_\rho \subset \zeta_\infty \}$$

denote the set of strategies (hypothetically) chosen between periods  $R$  and  $r$ . Let  $\kappa(\sigma_{-i} | \zeta_r) = |T_{1,r}|$  denote the number of times that strategy profile  $\sigma_{-i}$  is chosen before round  $r$  (along learning path  $\zeta_\infty$ ). Suppose for contradiction that there exists a strategy  $\sigma'_i \in \Sigma_i$  such that for each  $r > R$  and each



$\sigma_{-i} \in T_{R,r}$  it holds  $U_i(\sigma'_i, \sigma_{-i}) > U_i(\hat{\sigma}(\zeta_r), \sigma_{-i}) + \epsilon$ . Then since  $\kappa(\sigma_{-i} | \zeta_r)/r \rightarrow 0$  for each  $\sigma_{-i} \notin T_{R,r}$  there exists  $r_1^*$  such that for each  $r > r_1^*$  it must hold that

$$\begin{aligned} & \sum_{\sigma_{-i} \in \Sigma_{-i}} \frac{\kappa(\sigma_{-i} | \zeta_r)}{4r} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \sigma'_i(s_i, h_i) [s_i \varphi(h_i | 1, \sigma_{-i}) - (1-s_i) \varphi(h_i | 0, \sigma_{-i})] \\ & > \sum_{\sigma_{-i} \in \Sigma_{-i}} \frac{\kappa(\sigma_{-i} | \zeta_r)}{4r} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \hat{\sigma}_i(s_i, h_i | \zeta_r) [s_i \varphi(h_i | 1, \sigma_{-i}) - (1-s_i) \varphi(h_i | 0, \sigma_{-i})] + \frac{\epsilon}{2} \end{aligned} \quad (13)$$

where  $\varphi(h_i | \theta, \sigma_{-i}) = \Pr(h_i | \tilde{\theta} = \theta, \sigma_{-i})$ . Let

$$\kappa(h_i, \theta, \sigma_{-i} | \zeta_r) = |\{1 \leq \rho < r : \{h_i, \theta\} \in y(\rho) \text{ and } \sigma_{-i} = \hat{\sigma}_{-i}(\zeta_\rho)\}|.$$

By the *strong law of large numbers for conditional expectation (SSLNCE)* (see for instance Walk, 2008)

$$\frac{\kappa(h_i, \theta, \sigma_{-i} | \zeta_r)}{\sum_{h'_i \in H_i} \kappa(h'_i, \theta, \sigma_{-i} | \zeta_r)} \rightarrow \Pr(h_i | \tilde{\theta} = 1, \sigma_{-i})$$

almost surely as  $r \rightarrow \infty$  for each  $h_i \in H_i$  and each  $\theta \in \Theta$ . Notice that this relies crucially on the fact that in a given repetition a player's *strategy* cannot be correlated with the state of Nature or the player's private belief drawn *in this repetition*. Furthermore the SSLNCE holds even with partial and/or incomplete feedback and heterogeneous private belief support taking appropriate subsequences since each player will infinitely often decide at position  $n$  and given a history and private belief which will incline her to invest. Therefore for almost any  $\zeta_\infty$  and any  $\delta > 0$  there exists  $r^*$  such that

$$\left| \frac{\kappa(h_i, \theta, \sigma_{-i} | \zeta_r)}{\sum_{h'_i \in H_i} \kappa(h'_i, \theta, \sigma_{-i} | \zeta_r)} - \Pr(h_i | \tilde{\theta} = 1, \sigma_{-i}) \right| < \delta$$

for each  $r > r^*$ . Hence for the particular learning path chosen there exists  $r_2^* > r_1^*$  such that

$$\begin{aligned} & \sum_{\sigma_{-i} \in \Sigma_{-i}} \frac{\kappa(\sigma_{-i} | \zeta_r)}{4r} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \hat{\sigma}_i(s_i, h_i | \zeta_r) [s_i \varphi(h_i | 1, \sigma_{-i}) - (1-s_i) \varphi(h_i | 0, \sigma_{-i})] + \frac{\epsilon}{2} \\ & > \sum_{\sigma_{-i} \in \Sigma_{-i}} \frac{\kappa(\sigma_{-i} | \zeta_r)}{4r} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \hat{\sigma}_i(s_i, h_i | \zeta_r) \left[ s_i \left( \frac{\kappa(h_i, 1, \sigma_{-i})}{\sum_{h'_i \in H_i} \kappa(h'_i, 1, \sigma_{-i})} - \delta \right) \right. \\ & \quad \left. - (1-s_i) \left( \frac{\kappa(h_i, 0, \sigma_{-i})}{\sum_{h'_i \in H_i} \kappa(h'_i, 0, \sigma_{-i})} + \delta \right) \right] + \frac{\epsilon}{2} \end{aligned} \quad (14)$$

$$\begin{aligned} & = \frac{1}{4} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \hat{\sigma}_i(s_i, h_i | \zeta_r) \left[ s_i \sum_{\sigma_{-i} \in \Sigma_{-i}} \frac{\kappa(\sigma_{-i} | \zeta_r)}{r} \frac{\kappa(h_i, 1, \sigma_{-i} | \zeta_r)}{\sum_{h'_i \in H_i} \kappa(h'_i, 1, \sigma_{-i} | \zeta_r)} \right. \\ & \quad \left. - (1-s_i) \sum_{\sigma_{-i} \in \Sigma_{-i}} \frac{\kappa(\sigma_{-i} | \zeta_r)}{r} \frac{\kappa(h_i, 0, \sigma_{-i} | \zeta_r)}{\sum_{h'_i \in H_i} \kappa(h'_i, 0, \sigma_{-i} | \zeta_r)} - \delta \right] + \frac{\epsilon}{2} \end{aligned} \quad (15)$$

for each  $r > r_2^*$ . Now asymptotically for each  $\theta \in \Theta$  and each  $\sigma_{-i} \in T_{R,r}$  for  $r$  sufficiently large,  $\kappa(\sigma_{-i} | \zeta_r) / r = \kappa(\sigma_{-i}, \theta | \zeta_r) / \kappa(\theta | \zeta_r)$  due to the independence of the strategy chosen in a round from the drawn private belief in that round. Furthermore it holds  $\sum_{\sigma_{-i} \in \Sigma_{-i}} \kappa(h_i, \theta, \sigma_{-i} | \zeta_r) = \kappa(h_i, \theta | \zeta_r)$ . Third  $\sum_{h'_i \in H_i} \kappa(h'_i, \theta | \zeta_r) \approx \kappa(\theta | \zeta_r)$  respectively  $\sum_{h'_i \in H_i} \kappa(h'_i, \theta, \sigma_{-i} | \zeta_r) \approx \kappa(\theta, \sigma_{-i} | \zeta_r)$  for  $r$  sufficiently large taking appropriate subsequences. Finally recall that  $\hat{\varphi}_i(h_i | \theta; \zeta_r) = \kappa(h_i, \theta | \zeta_r) / \sum_{h'_i \in H_i} \kappa(h'_i, \theta | \zeta_r)$ . In summary there exists  $r_3^* > r_2^*$  such that for each  $r > r_3^*$

$$\begin{aligned}
&= \frac{1}{4} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \hat{\sigma}_i(s_i, h_i | \zeta_r) \left[ s_i \sum_{\sigma_{-i} \in \Sigma_{-i}} \frac{\kappa(\sigma_{-i} | \zeta_r)}{r} \frac{\kappa(h_i, 1, \sigma_{-i} | \zeta_r)}{\sum_{h'_i \in H_i} \kappa(h'_i, 1, \sigma_{-i} | \zeta_r)} \right. \\
&\quad \left. - (1 - s_i) \sum_{\sigma_{-i} \in \Sigma_{-i}} \frac{\kappa(\sigma_{-i} | \zeta_r)}{r} \frac{\kappa(h_i, 0, \sigma_{-i} | \zeta_r)}{\sum_{h'_i \in H_i} \kappa(h'_i, 0, \sigma_{-i} | \zeta_r)} - \delta \right] + \frac{\epsilon}{2} \\
&= \frac{1}{4} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \hat{\sigma}_i(s_i, h_i | \zeta_r) \left[ s_i \frac{\kappa(\sigma_{-i}, 1 | \zeta_r)}{\kappa(1 | \zeta_r)} \frac{\sum_{\sigma_{-i} \in \Sigma_{-i}} \kappa(h_i, 1, \sigma_{-i} | \zeta_r)}{\kappa(1, \sigma_{-i} | \zeta_r)} \right. \\
&\quad \left. - (1 - s_i) \frac{\kappa(0, \sigma_{-i} | \zeta_r)}{\kappa(0 | \zeta_r)} \frac{\sum_{\sigma_{-i} \in \Sigma_{-i}} \kappa(h_i, 0, \sigma_{-i} | \zeta_r)}{\kappa(0, \sigma_{-i} | \zeta_r)} - \delta \right] + \frac{\epsilon}{2} \tag{16}
\end{aligned}$$

$$= \frac{1}{4} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \hat{\sigma}_i(s_i, h_i | \zeta_r) [s_i \hat{\varphi}_i(h_i | 1; \zeta_r) - (1 - s_i) \hat{\varphi}_i(h_i | 0; \zeta_r) - \delta] + \frac{\epsilon}{2} \tag{17}$$

Note that  $\hat{\sigma}(\zeta_r)$  best responds to  $(s_i, h_i)$  at each  $h_i$  occurring a non-vanishing fraction of the time along path  $\zeta_\infty$ . Hence, if  $\sigma'_i \neq \hat{\sigma}(\zeta_r)$  for  $r > r_3^*$  there must exist a pair  $(s_i, h_i)$  at which  $\sigma'_i$  is worse than  $\hat{\sigma}_i(\zeta_r)$  in expected terms. Hence

$$\begin{aligned}
&\frac{1}{4} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \hat{\sigma}_i(s_i, h_i | \zeta_r) [s_i \hat{\varphi}_i(h_i | 1; \zeta_r) - (1 - s_i) \hat{\varphi}_i(h_i | 0; \zeta_r) - \delta] + \frac{\epsilon}{2} \\
&> \frac{1}{4} \sum_{h_i \in H_i} \int_{\underline{b}}^{\bar{b}} \frac{dG_1(s_i)}{s_i} \sigma'_i(s_i, h_i) [s_i \hat{\varphi}_i(h_i | 1; \zeta_r) - (1 - s_i) \hat{\varphi}_i(h_i | 0; \zeta_r) - \delta] + \frac{\epsilon}{2} \tag{18}
\end{aligned}$$

for  $r > r_3^*$ . Combining equations (13) – (18) and noting that  $\delta > 0$  has been chosen arbitrarily yields the desired contradiction.  $\square$

**Lemma D.1** (Remainder of the proof of Proposition 2). *Let  $\hat{\sigma}$  denote a strategy with associated assessments  $\hat{\varphi}$  (i.e.  $\hat{\varphi}$  is derived from  $\hat{\sigma}$  and  $\hat{\sigma}$  best responds to  $\hat{\varphi}$ ) and let  $(\varphi^*, \sigma^*)$  denote the rational assessment and strategy.*

*For each  $i = 1, \dots, n$  and each  $\epsilon > 0$  there exists  $\delta > 0$  such that  $\hat{\sigma}_i$  plays  $\epsilon$ -like  $\sigma_i^*$  at any  $h_i \in H_i$  such that  $\varphi_i^*(h_i | \theta) > 0$  for each  $\theta \in \Theta$  if  $\hat{\varphi}_i$  is  $\delta$ -close to  $\varphi_i^*$ .*

*Conversely for each  $i = 1, \dots, n$  and each  $\epsilon > 0$  there exists  $\delta > 0$  such that  $\hat{\varphi}_i$  is  $\epsilon$ -close to  $\varphi_i^*$  if for each  $j < i$   $\hat{\sigma}_j$  plays  $\delta$ -like  $\sigma_j^*$  at each  $h_j$  such that  $\varphi_j^*(h_j | \theta) > 0$  for each  $\theta \in \Theta$ .*

*Proof.* For the first statement let  $\varphi_i^*(h_i | \theta) > 0$  for each  $\theta \in \Theta$ . Recall that  $\hat{\sigma}_i$  plays  $\epsilon$ -like  $\sigma_i^*$  at  $h_i$  if there exists a set of private beliefs  $B_\epsilon$  of (unconditional) probability mass at least  $1 - \epsilon$  such that for each  $s_i \in B_\epsilon$  it holds  $|\hat{\sigma}_i(s_i, h_i) - \sigma_i^*(s_i, h_i)| < \epsilon$ . Since both  $\hat{\sigma}_i$  and  $\sigma_i^*$  are cutoff strategies the holds provided the respective cutoffs  $\hat{b}(\emptyset, h_i)$  and  $b^*(\emptyset, h_i)$  span a private belief interval of (unconditional) probability mass at most  $\epsilon$ . Notice that since we rule out ties, the set  $\{b^*(\emptyset, h_i)\}$  has mass zero. Accordingly there exists  $\epsilon_1 > 0$  such that the interval spanned by the cutoffs has probability mass at most  $\epsilon$  if  $|\hat{b}(\emptyset, h_i) - b^*(\emptyset, h_i)| < \epsilon_1$ . Cutoffs are respectively given by

$$\hat{b}(\emptyset, h_i) = \frac{\hat{\varphi}_i(h_i | 0)}{\hat{\varphi}_i(h_i | 1) + \hat{\varphi}_i(h_i | 0)} \quad \text{and} \quad b^*(\emptyset, h_i) = \frac{\varphi_i^*(h_i | 0)}{\varphi_i^*(h_i | 1) + \varphi_i^*(h_i | 0)}.$$

Assume that  $|\hat{\varphi}_i(h_i | \theta)| < \delta < \min\{\varphi_i^*(h_i | 1), \varphi_i^*(h_i | 0)\}$  for each  $\theta \in \Theta$ . Hence  $\hat{\varphi}_i(h_i | \theta) > 0$  for each  $\theta \in \Theta$  and furthermore

$$\begin{aligned} |\hat{b}(\emptyset, h_i) - b^*(\emptyset, h_i)| &= \left| \frac{\hat{\varphi}_i(h_i | 0)}{\hat{\varphi}_i(h_i | 1) + \hat{\varphi}_i(h_i | 0)} - \frac{\varphi_i^*(h_i | 0)}{\varphi_i^*(h_i | 1) + \varphi_i^*(h_i | 0)} \right| \\ &\leq \frac{\hat{\varphi}_i(h_i | 0) |\varphi_i^*(h_i | 1) - \hat{\varphi}_i(h_i | 1)| + \hat{\varphi}_i(h_i | 1) |\hat{\varphi}_i(h_i | 0) - \varphi_i^*(h_i | 0)|}{[\hat{\varphi}_i(h_i | 1) + \hat{\varphi}_i(h_i | 0)] + [\varphi_i^*(h_i | 1) + \varphi_i^*(h_i | 0)]} \\ &< \frac{\delta}{\varphi_i^*(h_i | 1) + \varphi_i^*(h_i | 0)}. \end{aligned}$$

Accordingly since  $\varphi_i^*(h_i | \theta) > 0$  for each  $\theta \in \Theta$ ,  $\hat{\sigma}_i$  plays  $\epsilon$ -like  $\sigma_i^*$  at  $h_i$  if  $\hat{\varphi}_i$  is  $\delta$ -close to  $\varphi_i^*$  for any  $\delta < \min\{\varphi_i^*(h_i | 1), \varphi_i^*(h_i | 0), \epsilon_1 * [\varphi_i^*(h_i | 1) + \varphi_i^*(h_i | 0)]\}$ .

For the conversion recall first that

$$|\hat{\varphi}_i(h_i | \theta) - \varphi_i^*(h_i | \theta)| = \left| \prod_{j < i} \int_{\underline{b}}^{\bar{b}} \hat{\sigma}_j(a_j | s_j, h_j) dG_\theta(s_j) - \prod_{j < i} \int_{\underline{b}}^{\bar{b}} \sigma_j^*(a_j | s_j, h_j) dG_\theta(s_j) \right|.$$

We distinguish two cases of histories. First assume that  $\varphi_i^*(h_i | \theta) > 0$  for each  $\theta \in \Theta$ . Accordingly for each  $h_j \subset h_i$ ,  $j < i$ , and each  $\theta \in \Theta$ ,  $\varphi_j^*(h_j | \theta) > 0$ . We proceed via induction. For  $h_1$  by definition  $\hat{\varphi}_1(h_1 | \theta) = \varphi_1^*(h_1 | \theta) = 1$  for each  $\theta \in \Theta$ . Hence, assume that  $\hat{\varphi}_j$  is  $\epsilon_1/2$ -close to  $\varphi_j^*$  for each  $j < i$  provided  $\hat{\sigma}_k$  plays  $\delta_k$ -like  $\sigma_k^*$  for each  $k < j$  at each  $h_k \subset h_i$ . Let  $a_k = \int_{\underline{b}}^{\bar{b}} \hat{\sigma}_k(a_k | s_k, h_k) dG_\theta(s_k)$  and  $b_k = \int_{\underline{b}}^{\bar{b}} \sigma_k^*(a_k | s_k, h_k) dG_\theta(s_k)$  for  $a_k \in h_i$  and  $h_k \subset h_i$ . Then

$$|\hat{\varphi}_i(h_i | \theta) - \varphi_i^*(h_i | \theta)| = \left| \prod_{k < i} a_k - \prod_{k < i} b_k \right| \leq a_{i-1} \left| \prod_{k < i-1} a_k - \prod_{k < i-1} b_k \right| + |a_{i-1} - b_{i-1}| \prod_{k < i-1} b_k.$$

By induction assumption  $\left| \prod_{k < i-1} a_k - \prod_{k < i-1} b_k \right| < \epsilon_1/2$ . Furthermore there exists  $\delta_{i-1}$  such that if  $\hat{\sigma}_{i-1}$  plays  $\delta_{i-1}$ -close to  $\sigma_{i-1}^*$  at  $h_{i-1} \subset h_i$  we can write

$$|a_{i-1} - b_{i-1}| \leq \delta_{i-1} G_\theta(B_{\delta_{i-1}}) + \delta_{i-1} \leq 2\delta_{i-1} < \epsilon_1/2$$

Finally for  $\delta_{i-1}$  sufficiently small  $0 < a_{i-1} < 1$ . Hence, we can choose  $\delta = \min_{j < i} \delta_j$  to obtain the desired result. On the other hand if  $\varphi_i^*(h_i | \theta) = 0$  let  $j_0$  be the maximal  $j < i$  such that  $\varphi_j^*(h_j | \theta) > 0$  for  $h_j \subset h_i$ . Then for each  $k \leq j_0$ ,  $\varphi_k^*(h_k | \theta) > 0$  and by the result above we can find  $\delta > 0$  such that if  $\hat{\sigma}_k$  plays  $\delta$ -close to  $\sigma_k^*$  for each  $k \leq j_0$  at each  $h_k \subseteq h_{j_0}$  then  $\hat{\varphi}_{j_0+1}(h_{j_0+1} | \theta)$  is  $\epsilon$ -close to  $\varphi_{j_0+1}^*(h_{j_0+1} | \theta)$ . Furthermore since  $\varphi_{j_0+1}(h_{j_0+1} | \theta) = 0$  for each  $\theta \in \Theta$  (since  $j_0$  was maximal) it must be that  $\hat{\varphi}_{j_0+1}(h_{j_0+1} | \theta) < \epsilon$ . Finally, since  $\hat{\varphi}_i(h_i | \theta) \leq \hat{\varphi}_{j_0+1}(h_{j_0+1} | \theta)$  the result follows for such histories as well.  $\square$

**Proof of Proposition 3.** Fix some ABEE  $\sigma$ . We start by deriving the associated analogy-based expectations  $\bar{\sigma}$ . Fix some player  $i$ , some  $h_j \in H_j$  and some  $\theta \in \Theta$ . For given setting  $k$  the probability of history  $h_j$  is straightforwardly defined via

$$\varphi^k(h_j | \theta) = \prod_{\ell < j} \sum_{s_\ell \in \text{supp}(dG_\theta^k)} \sigma_\ell(a_\ell | s_\ell, h_\ell) dG_\theta^k(s_\ell)$$

where  $a_\ell = h_j(\ell)$  and  $h_\ell \subset h_j$ . The probability measure  $\nu_\sigma$  is thus given by

$$\nu_\sigma(k, \theta, s_j, h_j) = \varphi^k(h_j | \theta) dG_\theta^k(s_j) \frac{\pi_k}{2}$$

and the associated conditional probabilities satisfy

$$\nu_\sigma(k, s_j | h_j, \theta) = \frac{\pi_k dG_\theta^k(s_j) \varphi^k(h_j | \theta)}{\sum_{k'=1}^K \pi_{k'} \varphi^{k'}(h_j | \theta) \int_{s \in \text{supp}(G_\theta^k)} dG_\theta^k(s)} = \frac{\pi_k \varphi^k(h_j | \theta)}{\sum_{k'=1}^K \pi_{k'} \varphi^{k'}(h_j | \theta)} dG_\theta^k(s_j).$$

Hence, for each  $i = 1, \dots, n$

$$\bar{\sigma}_i(h_j, \theta) = \sum_{k=1}^K \int_{s_j \in \text{supp}(G_\theta^k)} \frac{\pi_k \varphi^k(h_j | \theta)}{\sum_{k'=1}^K \pi_{k'} \varphi^{k'}(h_j | \theta)} dG_\theta^k(s_j) \sigma_j(s_j, h_j).$$

We now prove via induction that for each  $i$   $\bar{\varphi}$  satisfies  $\bar{\varphi}(h_i | \theta) = \sum_{k=1}^K \pi_k \varphi^k(h_i | \theta)$ . For  $i = 1$  this is clear since  $\varphi^k(h_1 | \theta) = 1$  and  $\sum_{k=1}^K \pi_k = 1$ . Assume that the claim holds for all  $j \leq i$ . For  $h_i \in H_i$  and  $a_i \in A$  we obtain

$$\begin{aligned} \bar{\varphi}((h_i, a_i) | \theta) &= \bar{\varphi}(h_i | \theta) * \bar{\sigma}_i(a_i | h_i, \theta) \\ &= \left[ \sum_{k_1=1}^K \pi_{k_1} \varphi^{k_1}(h_i | \theta) \right] * \left[ \sum_{k_2=1}^K \int_{s_i \in \text{supp}(G_\theta^{k_2})} \frac{\pi_{k_2} \varphi^{k_2}(h_i | \theta)}{\sum_{k_3=1}^K \pi_{k_3} \varphi^{k_3}(h_i | \theta)} dG_\theta^{k_2}(s_i) \sigma_i(a_i | s_i, h_i) \right] \\ &= \sum_{k_2=1}^K \pi_{k_2} \varphi^{k_2}(h_i | \theta) \int_{s_i \in \text{supp}(G_\theta^{k_2})} \sigma_i(a_i | s_i, h_i) dG_\theta^{k_2}(s_i) \\ &= \sum_{k_2=1}^K \pi_{k_2} \varphi^{k_2}((h_i, a_i) | \theta). \end{aligned}$$

For the purpose of deriving the iteratively undominated strategies the ex-ante expected utility for player  $i$  given analogy-based expectations  $\bar{\sigma}$  is given by

$$\begin{aligned} U_i(\sigma_i, \bar{\sigma}) &= \sum_{k=1}^K \pi_k \sum_{h_i \in H_i} \int_{s_i \in \text{supp}(G_1^k)} \sigma_i(s_i, h_i) [s_i \bar{\varphi}(h_i | 1) - (1 - s_i) \bar{\varphi}(h_i | 0)] \frac{dG_1^k(s_i)}{s_i} \\ &= \sum_{h_i \in H_i} \int_{s_i \in \bigcup_k \text{supp}(G_1^k)} \sigma_i(s_i, h_i) [s_i \bar{\varphi}(h_i | 1) - (1 - s_i) \bar{\varphi}(h_i | 0)] \frac{\sum_{k=1}^K dG_1^k(s_i)}{s_i}. \end{aligned}$$

Clearly, the optimal response to assessments  $\bar{\varphi}$  requires  $\sigma_i(s_i, h_i) = 1$  if  $s_i \bar{\varphi}(h_i | 1) > (1 - s_i) \bar{\varphi}(h_i | 0)$  and  $\sigma_i(s_i, h_i) = 0$  if  $s_i \bar{\varphi}(h_i | 1) < (1 - s_i) \bar{\varphi}(h_i | 0)$ . Straightforward manipulations show that this is equivalent to sequential best response. Moreover assessments  $\bar{\varphi}(h_i | \theta)$  do only depend on strategies for periods  $j < i$ . Therefore the strategy for period 1 does not depend on assessments and is uniquely

defined except at  $s_i = 1/2$  which occurs with strictly positive probability only in non-generic settings. Therefore assessments for period 2 are uniquely defined yielding a uniquely defined strategy in period 2 (again outside non-generic settings). This argument is straightforwardly extended to all periods proving the claim.  $\square$

**Proof of Proposition 4.** The proof invokes the standard inductive argument: In the first period assessments trivially coincide with analogy-based assessments and strategic responses are uniquely defined by myopic Bayesian rationality in each repetition and coincide with both ABEE and rational play (since assessments are trivial). Assume the claims hold for periods  $j < i$ . For some learning path  $\zeta_r$  and some settings  $k$  define frequencies  $\kappa(k, h_i, \theta | \zeta_r)$  in the obvious way. Assessments for period  $i$  may then be decomposed as

$$\begin{aligned}\hat{\varphi}(h_i | \theta; \zeta_r) &= \frac{\kappa(h_i, \theta | \zeta_r)}{\sum_{h'_i} \kappa(h'_i, \theta | \zeta_r)} = \sum_{k=1}^K \frac{\kappa(k, h_i, \theta | \zeta_r)}{\sum_{k'} \sum_{h'_i} \kappa(k', h'_i, \theta | \zeta_r)} * \frac{\sum_{h'_i} \kappa(k, h'_i, \theta | \zeta_r)}{\sum_{h'_i} \kappa(k, h'_i, \theta | \zeta_r)} \\ &= \sum_{k=1}^K \frac{\kappa(k, h_i, \theta | \zeta_r)}{\sum_{h'_i} \kappa(k, h'_i, \theta | \zeta_r)} * \frac{\sum_{h'_i} \kappa(k, h'_i, \theta | \zeta_r)}{\sum_{k'} \sum_{h'_i} \kappa(k', h'_i, \theta | \zeta_r)}.\end{aligned}$$

By induction assumption, an adaption of the second part of Lemma D.1 and the SLLN for conditional expectations (SLLNCE) the first part of each summand converges to  $\varphi^k(h_i | \theta)$  for each  $k = 1, \dots, K$  along almost any  $\zeta_\infty$ . On the other hand the SLLNCE and independent draws of setting and state of Nature imply that the second part converges to  $\pi_k$  for each  $k$  along almost any  $\zeta_\infty$ . In conjunction with Proposition 3 this implies that  $\hat{\varphi}(h_i | \theta; \zeta_r)$  converges to  $\bar{\varphi}(h_i | \theta)$  along almost any  $\zeta_\infty$  for each  $h_i \in H_i$  and each  $\theta \in \Theta$ . Hence, assessments almost surely eventually become arbitrarily close to analogy-based assessments in period  $i$  as well. Induction and proof are finished by adapting the first part of Lemma D.1.  $\square$

**Proof of the Corollary.** We provide a generic example: For  $K = 2$  let  $G_\theta^k$ ,  $k = 1, 2$  be continuous and satisfy  $G_0^1(1/2)/G_1^1(1/2) < G_0^2(1/2)/G_1^2(1/2)$  and let  $\pi_1 = 1 - \pi_2 = \pi$  where  $0 < \pi < 1$ .

After a rejection in the first period ( $a_1 = 0$ ), second period's assessments are given by  $\bar{\varphi}(0 | \theta) = \pi G_\theta^1(1/2) + (1 - \pi) G_\theta^2(1/2)$  and  $\varphi^{*k}(0 | \theta) = G_\theta^k(1/2)$  for  $k = 1, 2$ . Accordingly public beliefs satisfy

$$\begin{aligned}b^{2*}(\emptyset, (0)) &= \frac{G_1^2(1/2)}{G_1^2(1/2) + G_0^2(1/2)} \\ &< \frac{\pi G_1^1(1/2) + (1 - \pi) G_1^2(1/2)}{\pi (G_1^1(1/2) + G_0^1(1/2)) + (1 - \pi) (G_1^2(1/2) + G_0^2(1/2))} = \bar{b}_2(\emptyset, (0)) \\ &< \frac{G_1^1(1/2)}{G_1^1(1/2) + G_0^1(1/2)} = b^{1*}(\emptyset, (0)).\end{aligned}$$

Accordingly, in ABEE by continuity of private belief distributions players eventually imitate the first player's rejection too often in environment  $E_1$  and too seldom in environment  $E_2$ .  $\square$

**Proof of Lemma 2.** Fix period  $i$ . Given the vector of assessments  $(\varphi^1, \dots, \varphi^K)$  the expected payoff to strategy  $\sigma$  is given by

$$\begin{aligned} U_i(\sigma | \varphi^1, \dots, \varphi^K) &= \frac{1}{4} \sum_{k=1}^K \pi_k \sum_{h_i \in H_i} \int_0^1 \sigma(s, h_i) [\varphi^k(h_i | 1) dG_1^k(s) - \varphi^k(h_i | 0) dG_0^k(s)] \\ &= \frac{1}{4} \sum_{h_i \in H_i} \int_0^1 \sigma(s, h_i) \sum_{k=1}^K \pi_k [\varphi^k(h_i | 1) dG_1^k(s) - \varphi^k(h_i | 0) dG_0^k(s)]. \end{aligned}$$

Obviously, this expression is maximized on  $\Sigma$  by selecting  $\sigma(s, h_i) = 1$  whenever  $\bar{U}(s, h_i) = \sum_{k=1}^K \pi_k [\varphi^k(h_i | 1) dG_1^k(s) - \varphi^k(h_i | 0) dG_0^k(s)] > 0$  and  $\sigma(s, h_i) = 0$  whenever  $\bar{U}(s, h_i) < 0$ . This straightforwardly yields the optimal strategy  $\sigma^*$ .  $\square$

**Proof of Proposition 5.** We provide a generic example with  $K = 2$ . For  $k = 1, 2$ ,  $G_\theta^k$  is continuously distributed on  $\text{supp}(G_\theta^k) = [1 - a_k, a_k]$  according to conditional densities  $g_0^k(s) = 2(1-s)/(2a_k - 1)$  and  $g_1^k(s) = 2s/(2a_k - 1)$ . W.l.o.g.  $a_1 < a_2$ . The benchmark strategy  $\sigma^*$  can be straightforwardly derived as  $\sigma^*(s, h) = 1$  if  $s > 1 - \check{b}(\emptyset, h | \varphi^1, \dots, \varphi^K)$  and  $\sigma^*(s, h) = 0$  if  $s < 1 - \check{b}(\emptyset, h | \varphi^1, \dots, \varphi^K)$  where the *benchmark public belief* is given by

$$\check{b}(\emptyset, h | \varphi^1, \dots, \varphi^K) = \frac{\sum_{k=1}^K \frac{\pi_k}{2a_k - 1} \varphi^k(h | 1)}{\sum_{k=1}^K \frac{\pi_k}{2a_k - 1} [\varphi^k(h | 1) + \varphi^k(h | 0)]}.$$

On the other hand in the ABEE players invest if  $s > 1 - b(\emptyset, h | \varphi^1, \dots, \varphi^K)$  and reject if  $s < 1 - b(\emptyset, h | \varphi^1, \dots, \varphi^K)$  where the *ABEE public belief* is given by

$$b(\emptyset, h | \varphi^1, \dots, \varphi^K) = \frac{\sum_{k=1}^K \pi_k \varphi^k(h | 1)}{\sum_{k=1}^K \pi_k [\varphi^k(h | 1) + \varphi^k(h | 0)]}.$$

As before continuity of private beliefs implies that ABEE and benchmark strategy differ unless at all histories occurring with strictly positive probability ABEE and benchmark public belief either coincide or lie both strictly within the same cascade set. By straightforward algebraic transformation we obtain

$$b(\emptyset, h | \varphi^1, \dots, \varphi^K) - \check{b}(\emptyset, h | \varphi^1, \dots, \varphi^K) = \frac{2(a_2 - a_1) [\varphi^2(h | 1) \varphi^1(h | 0) - \varphi^1(h | 1) \varphi^2(h | 0)]}{(2a_2 - 1)(2a_1 - 1)}.$$

Therefore ABEE and benchmark strategy coincide if and only if for each history occurring a strictly positive fraction of the time either  $\varphi^1(h | 1) * \varphi^2(h | 0) = \varphi^2(h | 1) * \varphi^1(h | 0)$ , or ABEE and benchmark public belief lie strictly within a cascade set. Generically this is not satisfied. For instance in period 2

$$\varphi^k(a_1 = 1 | \theta) = 1 - G_\theta^k(1/2) = \begin{cases} \frac{1}{4} + \frac{a_k}{2} & \text{if } \theta = 1 \\ \frac{1}{4} + \frac{1-a_k}{2} & \text{if } \theta = 0 \end{cases}$$

which implies that  $b(\emptyset, (1) | \varphi^1, \varphi^2) - \check{b}(\emptyset, (1) | \varphi^1, \varphi^2) = \frac{(a_2 - a_1)^2}{(2a_2 - 1)(2a_1 - 1)} > 0$  and  $b(\emptyset, (1) | \varphi^1, \varphi^2) = \frac{1}{4} + \frac{\pi a_1 + (1-\pi)a_2}{2} \in (1 - a_1, a_2)$  since  $a_1 < \pi a_1 + (1-\pi)a_2 < a_2$  and  $a_2 > a_1 > 1/2$ . A similar result holds for  $h_2 = (0)$ .  $\square$

**Proof of Proposition 6.** We calculate the second derivative of

$$\check{b}(s, h, \epsilon) = \frac{b(s, \emptyset) [b(\emptyset, h) + \epsilon]}{b(s, \emptyset) [b(\emptyset, h) + \epsilon] + [1 - b(s, \emptyset)] [1 - b(\emptyset, h) - \epsilon]}$$

with respect to  $\epsilon$ . We obtain

$$\frac{\partial^2 \check{b}(s, h, \epsilon)}{\partial \epsilon^2} = \frac{2 b(s, \emptyset) [1 - b(s, \emptyset)] [1 - 2 b(s, \emptyset)]}{[b(s, \emptyset) [b(\emptyset, h) + \epsilon] + [1 - b(s, \emptyset)] [1 - b(\emptyset, h) - \epsilon]]^3}.$$

Obviously,  $\partial^2 \check{b}(s, h, \epsilon)/\partial \epsilon^2 > 0$  if  $b(s, \emptyset) < 1/2$  and  $\partial^2 \check{b}(s, h, \epsilon)/\partial \epsilon^2 < 0$  if  $b(s, \emptyset) > 1/2$ . Therefore  $\check{b}(s, h, \epsilon)$  is strictly convex in  $\epsilon$  if  $b(s, \emptyset) < 1/2$  and strictly concave in  $\epsilon$  if  $b(s, \emptyset) > 1/2$ . Furthermore clearly,  $b'(s, h) = \check{b}(s, h, \tilde{\epsilon}_h)$ . Thus if  $\eta_h > 0$  by Jensen's inequality for  $b(s, \emptyset) < 1/2$ ,

$$E_{\tilde{\epsilon}_h} [b'(s, h)] > \check{b}(s, h, E_{\tilde{\epsilon}_h} [\tilde{\epsilon}_h]) = \check{b}(s, h, 0) = b(s, h)$$

and for  $b(s, \emptyset) > 1/2$

$$E_{\tilde{\epsilon}_h} [b'(s, h)] < \check{b}(s, h, E_{\tilde{\epsilon}_h} [\tilde{\epsilon}_h]) = \check{b}(s, h, 0) = b(s, h).$$

□

**Proof of Lemma 3.** Rewrite expected payoff in the noisy social learning game of strategy  $\sigma$  given probabilities  $Pr(h | \theta)$  as

$$U_{\boldsymbol{\eta}}(\sigma) = \frac{1}{4n} \int_{\underline{b}}^{\bar{b}} \int_0^1 \sigma(s, x) [f(x | 1) dG_1(s) - f(x | 0) dG_0(s)] dx$$

where

$$f(x | \theta) = \sum_{h \in H} Pr(h | \theta) \phi_h(x - b(\emptyset, h))$$

is the probability density function of the random variable  $b'(\emptyset, \tilde{h})$ . A similar separability argument as we have used before then implies that the optimal strategy  $\sigma_{\boldsymbol{\eta}}^* \in \Sigma_{\boldsymbol{\eta}}$  is given by

$$\sigma_{\boldsymbol{\eta}}^*(s, x) = \begin{cases} 1 & \text{if } b(s, \emptyset) > \frac{\sum_{h \in H} Pr(h|0) \phi_h(x - b(\emptyset, h))}{\sum_{h \in H} [Pr(h|1) + Pr(h|0)] \phi_h(x - b(\emptyset, h))} \\ 0 & \text{if } b(s, \emptyset) < \frac{\sum_{h \in H} Pr(h|0) \phi_h(x - b(\emptyset, h))}{\sum_{h \in H} [Pr(h|1) + Pr(h|0)] \phi_h(x - b(\emptyset, h))} \end{cases}.$$

□

**Proof of Proposition 7.** We show that generically the first derivative of

$$t(x | \boldsymbol{\eta}) = \frac{\sum_{h \in H} Pr(h | 0) \phi_h(x - b(\emptyset, h))}{\sum_{h \in H} [Pr(h | 1) + Pr(h | 0)] \phi_h(x - b(\emptyset, h))}$$

with respect to  $\eta_h^2$  is different from zero provided  $\eta_{h'}$  is sufficiently large for some  $h' \in H$ . This in turn implies that for generic  $\boldsymbol{\eta}$ ,  $t(x | \boldsymbol{\eta}) \neq 1 - x$ .

As a first step we calculate

$$\begin{aligned}
\frac{\partial \phi_h(x - b(\varnothing, h))}{\partial \eta_h^2} &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x - b(\varnothing, h))^2}{2\eta_h^2}\right) \left(-\frac{1}{2\eta_h^2 \sqrt{\eta_h^2}}\right) \\
&\quad + \frac{1}{\sqrt{2\pi \eta_h^2}} \exp\left(-\frac{(x - b(\varnothing, h))^2}{2\eta_h^2}\right) \frac{(x - b(\varnothing, h))^2}{2\eta_h^2 \eta_h^2} \\
&= \phi_h(x - b(\varnothing, h) \mid \varphi^*) \frac{(x - b(\varnothing, h))^2 - \eta_h^2}{2[\eta_h^2]^2}.
\end{aligned}$$

This implies that  $\partial \phi_h(x - b(\varnothing, h)) / \partial \eta_h^2 \neq 0$  whenever

$$x - b(\varnothing, h) \neq \eta_h \quad (19)$$

and  $\phi_h(x - b(\varnothing, h) \mid \varphi^*) > 0$  which is approximately satisfied if

$$|x - b(\varnothing, h) \mid \varphi^*| < 3\eta_h. \quad (20)$$

Denote by  $N = \sum_{h \in H} [Pr(h \mid 1) + Pr(h \mid 0)] \phi_h(x - b(\varnothing, h))$ . The derivative of  $t(x \mid \boldsymbol{\eta})$  with respect to  $\eta_h^2$  is then given by

$$\begin{aligned}
&\frac{\partial t(x \mid \boldsymbol{\eta})}{\partial \eta_h^2} \\
&= \frac{1}{N^2} \left\{ Pr(h \mid 0) \frac{\partial \phi_h(x - b(\varnothing, h))}{\partial \eta_h^2} * \sum_{h' \in H} [Pr(h' \mid 1) + Pr(h' \mid 0)] \phi_{h'}(x - b(\varnothing, h')) \right. \\
&\quad \left. - [Pr(h \mid 1) + Pr(h \mid 0)] \frac{\partial \phi_h(x - b(\varnothing, h))}{\partial \eta_h^2} \sum_{h' \in H} Pr(h' \mid 0) \phi_{h'}(x - b(\varnothing, h)) \right\} \\
&= \frac{1}{N^2} \frac{\partial \phi_h(x - b(\varnothing, h))}{\partial \eta_h^2} \sum_{h' \in H} \phi_{h'}(x - b(\varnothing, h)) [Pr(h' \mid 1) Pr(h \mid 0) - Pr(h \mid 1) Pr(h' \mid 0)].
\end{aligned}$$

From the argumentation above  $\partial t(x \mid \boldsymbol{\eta}) / \partial \eta_h^2 \neq 0$  can only hold for  $x$  satisfying (19) and (20). On the other hand the summand for  $h$  in the sum in the last line is zero. Hence, a necessary condition for the sum to be different from zero is existence of at least one  $h' \neq h$  such that  $\phi_{h'}(x - b(\varnothing, h')) \neq 0$  for at least one  $x$  satisfying (19) and (20). Since generically  $b(\varnothing, h') \neq b(\varnothing, h)$  this requires that  $\eta_{h'}$  is sufficiently large. Conversely, existence of such a history  $h'$  is generically also sufficient for the sum to be different from zero. To see this assume first that there exists exactly one such history  $h' \neq h$ . In this case  $\partial t(x \mid \boldsymbol{\eta}) / \partial \eta_h^2 \neq 0$  if  $Pr(h' \mid 1) Pr(h \mid 0) \neq Pr(h' \mid 0) Pr(h \mid 1)$ . Not being an identity, this is generically satisfied (see Smith and Sørensen, 2000, p.389). More generally if more than one  $h'$  satisfy the condition the sum is one of several generically non-zero terms each weighted by  $\phi_{h'}(x - b(\varnothing, h))$ . Clearly, this sum will generically be different from zero.  $\square$

**Benefits of Overweighting with Heterogeneous Preferences (Section 7.3).** We have  $K = 2$  and  $\alpha_1 = \alpha_2 = 1/2$ . The distribution of private beliefs in game  $k = 1, 2$ , is given by  $\text{supp}(G_\theta^k) = \{1 - q_k, q_k\}$  and  $Pr(b(\tilde{s}, \varnothing) = q_k \mid \tilde{\theta} = 1) = Pr(b(\tilde{s}, \varnothing) = 1 - q_k \mid \tilde{\theta} = 0) = q_k$  where  $1/2 < q_1 < q_2 < 1$ . Furthermore in game 2 only a fraction  $(1 - 2\xi_2)$  has standard preferences while there exist  $\xi_2$  noise players each which always invest or always reject respectively. In both environments in the first two periods



players optimally follow private information. (In game 1 we make this assumption for simplicity. It can be justified for instance by assuming a small positive fraction  $0 < \xi_1 \ll \xi_2$  of noise players in this game as well.) In the third period clearly imitating two similar decisions in period 3 is optimal in game 1 while in game 2 following private information is optimal provided  $q_2$  and  $\xi_2$  jointly satisfy

$$[\xi_2 + (1 - 2\xi_2)q_2]^2 (1 - q_2) < [\xi_2 + (1 - 2\xi_2)(1 - q_2)]^2 q_2 \Leftrightarrow \left( \frac{\xi_2}{1 - 2\xi_2} \right)^2 > q_2 (1 - q_2).$$

On the other hand in the mixed environment assessments in the third period are given by

$$\begin{aligned} \bar{\varphi}^*((1, 1) | \tilde{\theta} = 1) &= \bar{\varphi}^*((0, 0) | \tilde{\theta} = 0) = \left[ \frac{q_1}{2} + \frac{\xi_2}{2} + \frac{1 - 2\xi_2}{2} q_2 \right]^2, \\ \bar{\varphi}^*((1, 1) | \tilde{\theta} = 0) &= \bar{\varphi}^*((0, 0) | \tilde{\theta} = 1) = \left[ \frac{1 - q_1}{2} + \frac{\xi_2}{2} + \frac{1 - 2\xi_2}{2} (1 - q_2) \right]^2. \end{aligned}$$

Consider the heterogeneous updating rule model of March and Ziegelmeyer (2009). A player of type  $\beta \in (0, \infty)$  in game  $k$  imitates the first two decisions in period 3 iff  $\bar{\varphi}^*((1, 1) | 1) (1 - q_k)^\beta > \bar{\varphi}^*((1, 1) | 0) q_k^\beta$  and follows private information if the inequality is reversed. Accordingly provided

$$\frac{\log(\bar{\varphi}^*((1, 1) | 1)) - \log(\bar{\varphi}^*((1, 1) | 0))}{\log(q_2) - \log(1 - q_2)} < \beta < \frac{\log(\bar{\varphi}^*((1, 1) | 1)) - \log(\bar{\varphi}^*((1, 1) | 0))}{\log(q_1) - \log(1 - q_1)}$$

the player correctly imitates in game 1 and correctly follows private information in game 2. Since  $q_1 < q_2$  this interval is well-defined. Furthermore the lower bound strictly exceeds 1 provided

$$q_2 < \frac{q_1^2}{q_1^2 + (1 - q_1)^2}$$

and

$$\frac{\sqrt{q_2(1 - q_2)} - 2q_2(1 - q_2)}{(2q_2 - 1)^2} < \xi_2 < \frac{2\sqrt{q_2(1 - q_2)} - 2q_2(1 - q_2) - q_2(1 - q_1) - q_1(1 - q_2)}{(2q_2 - 1)^2}.$$

Hence, for these values of the parameters, overweighting is profitable with  $n = 3$ .

□